APPENDIX A
Elementary Statistics

The purpose of this appendix is to provide a quick and informative review of elementary statistics. Statistics is used in almost every facet of every life form aviation to weather prediction. In the field of finance, one of its primary uses is to characterize the rate of return distributions of risky securities or portfolios of securities, although the principles apply to prices changes, earnings, or cash flows of almost any sort. To clarify the use of the statistical concepts used in this appendix, numerous illustrations are provided. To make the concepts in the illustrations as usable as possible, we demonstrate how the computations the computations can be performed using Microsoft Excel add-ins.

OBJECTIVES

After reviewing this Appendix, you should be able to:

1. Understand the difference between a population and a sample.
2. Understand the statistical properties of a probability distribution.
3. Understand the properties of expectation operators.
4. Estimate properties of population from a sample of observations drawn from the distribution.
5. Understand the properties of important continuous distributions including the normal distribution, the chi-square distribution, the *t*-distribution, and the *F*-distribution.
6. Test the hypothesis that a given data series approximates the normal distribution.
7. Test the hypothesis that the mean of a population is zero.
8. Test the hypothesis that the means of two samples are equal when the samples have equal and unequal variances.
9. Test the hypothesis that the means in a paired sample are equal.
10. Understand the distinction between Type I and Type II errors in statistical inference.
11. Understand *p*-values and the power of tests.
12. Test the hypothesis that the variance of two samples are equal.

13. Test the hypothesis that a time series is autocorrelated.
14. Understand the relevance of the Central Limit Theorem in statistical inference.

POPULATION VERSUS SAMPLE

The need for statistics stems from a lack of complete information about a particular process. Statisticians refer to the total collection of observations or measurements from the process as the (finite- *or* infinite-sized) *population*. Data taken from the population via a particular study or experiment make up a (finite-sized) *sample*.

In practice, Greek letters are commonly used to denote quantities that characterize the population (such as μ or σ). These values are referred to as parameters and are generally considered to be fixed and unknown. Parameter estimates, denoted here by Greek letters with hats (such as $\hat{\mu}$ or $\hat{\sigma}$), are statistics calculated from the sample that are used as a best guess for the true parameter. Because we may never know the values of the true population parameters, we associate a value known a *standard error*, denoted $s_{\hat{\mu}}$, with each estimate. Thus in using statistical methods, we can obtain estimates for the relevant parameters and also quantify their uncertainty.

Summary of the Statistical Method

1. Identify the problem of interest.
2. Draw a random sample from the population.
3. Perform statistical tests on the sampled data.
4. Make inferences about the relevant population.

RANDOM VARIABLES

A *random variable* is a variable that takes on different values, each with a probability less than or equal to 1. The process that generates a random variable is called a *probability distribution*. It can be thought of as a list of all possible values of the variable and the probability that each will occur. A coin toss, for example, can be interpreted as a random variable generated from a *binomial probability distribution*.

A *discrete random variable* may take on only a specific number of real values. Consider the outcomes from rolling a pair of dice. The possible outcomes range from 2 to 12. If the dice are fair, each side of each die has an equal probability (i.e., a one in six chance) of appearing. If we enumerate all possible outcomes, a total of 2 can appear with only one combination—(1,1), a total of 3 can appear with two combinations—(1,2) and (2,1), a total of four can appear with three combinations—(1,3), (2,2), and (3,1), and so on. Figure A.1, Panel A shows the *frequency distribution* of possible outcomes. A value of 7 appears most frequently at $f_7 = 6$. The total number of possible outcomes is

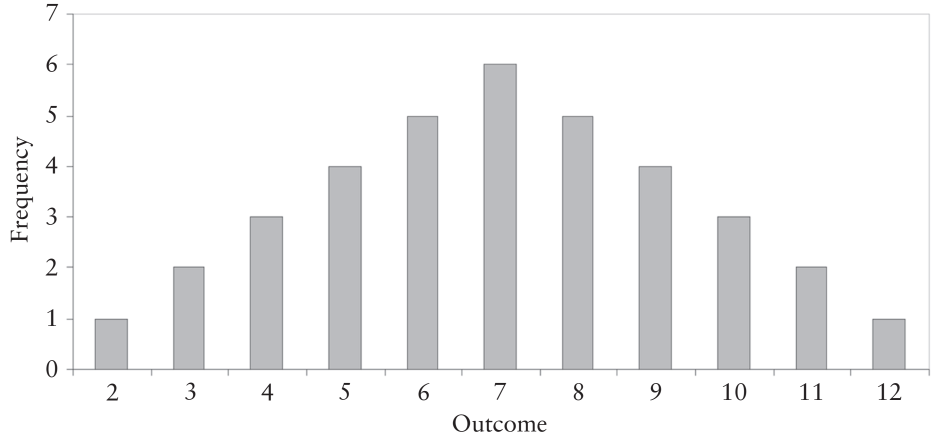
$$\sum_{i=2}^{12} f_i = 36$$

If we rescale the frequencies so they add up to one, that is,

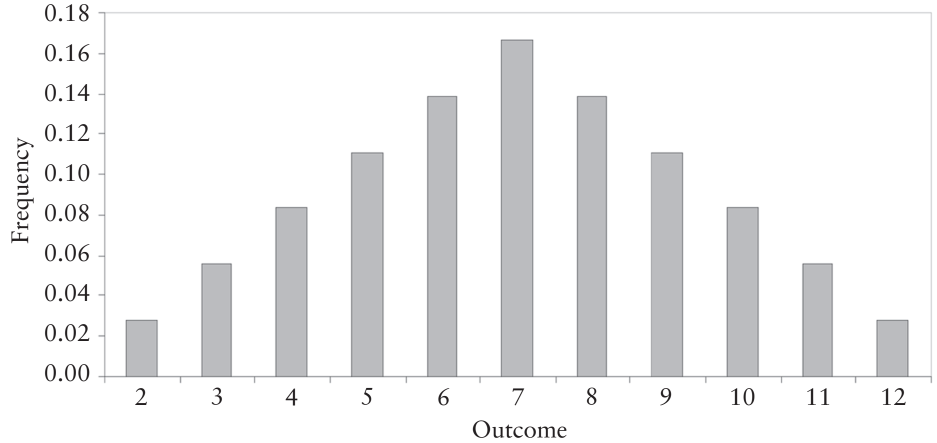
$$\sum_{i=2}^{12} \frac{f_i}{36} = \sum_{i=2}^{12} p_i = 1$$

we obtain the *discrete probability density function* (or discrete pdf) shown in Figure A.1, Panel B.

FIGURE A.1 Frequency and probability distributions of outcomes from rolling a pair of fair dice.
Panel A. Frequency distribution



Panel B. Probability distribution



Mean

A probability distribution is often characterized by its mean and variance.[1] The definitions of mean and variance, in turn, are defined in terms of the *expectations operator E*. Assume that $X_1, X_2, X_3, \ldots, X_N$ represent the N possible outcomes associated with the random variable X (i.e., the *population*). The *mean* or *expected value* of X, denoted μ_X, is defined as

$$\mu_X = E(X) = \sum_{i=1}^{N} p_i X_i \tag{A.1}$$

where p_i is the probability that X_i occurs, and the sum of the probabilities equals 1, that is,

$$\sum_{i=1}^{N} p_i = 1$$

Note that the mean is simply a weighted average of the possible outcomes, where the probabilities serve as outcome weights. Table A.1 shows the individual terms of the summation (A.1) for the above dice rolling illustration. The mean is 7. Note that μ_X is the mean of the population and is distinct from the *sample mean*, which is the average of the outcomes in a sample of size n (where $n < N$) drawn from the underlying distribution. The sample mean is denoted $\hat{\mu}_X$.

TABLE A.1 Mean and variance of outcomes from rolling a pair of fair dice.

Outcome, X_i	**Frequency,** f_i	**Probability,** p_i	**Expected Value,** p_iX_i	**Variance,** $p_i[X_i - E(X)]^2$
2	1	0.0278	0.0556	0.6944
3	2	0.0556	0.1667	0.8889
4	3	0.0833	0.3333	0.7500
5	4	0.1111	0.5556	0.4444
6	5	0.1389	0.8333	0.1389
7	6	0.1667	1.1667	0.0000
8	5	0.1389	1.1111	0.1389
9	4	0.1111	1.0000	0.4444
10	3	0.0833	0.8333	0.7500
11	2	0.0556	0.6111	0.8889
12	1	0.0278	0.3333	0.6944
Total	36	1.0000	7.0000	5.8333

[1] Indeed, under the capital asset pricing model discussed in Chapter 3, risky securities/portfolios are evaluated solely on the basis of these two parameters.

Variance and Standard Deviation

The *variance* of a random variable measures the dispersion of the distribution around the mean. The variance, denoted σ_X^2, is defined as

$$Var(X) = \sigma_X^2 = E[X - E(X)]^2 = \sum_{i=1}^{N} p_i[X_i - E(X)]^2 \tag{A.2}$$

Like the mean, the variance is a weighted average of the squares of the deviations of the outcomes on X from its expected value, with the probabilities serving as weights. Table A.1 also shows the individual terms of the summation (A-2) for the dice rolling illustration. The variance is 5.8333. The (positive) square root of the variance is called the *standard deviation*. The standard deviation (or variance) of a rate of return distribution is a commonly used measure of the total risk of a security.

Covariance and Correlation

In many applications in this book, we are interested in the *joint distribution* of X with a second random variable Y. With a joint distribution, the outcomes are in terms of both X and Y, and the probabilities are joint probabilities of the X-Y pair occurring. The *covariance* of X and Y, denoted σ_{XY}, is defined as

$$\begin{aligned} Cov(X, Y) &= \sigma_{XY} = E[(X - E(X))(Y - E(Y))] \\ &= \sum_{i=1}^{N} \sum_{j=1}^{N} p_{ij}(X_i - E(X))(Y_j - E(Y)) \end{aligned} \tag{A.3}$$

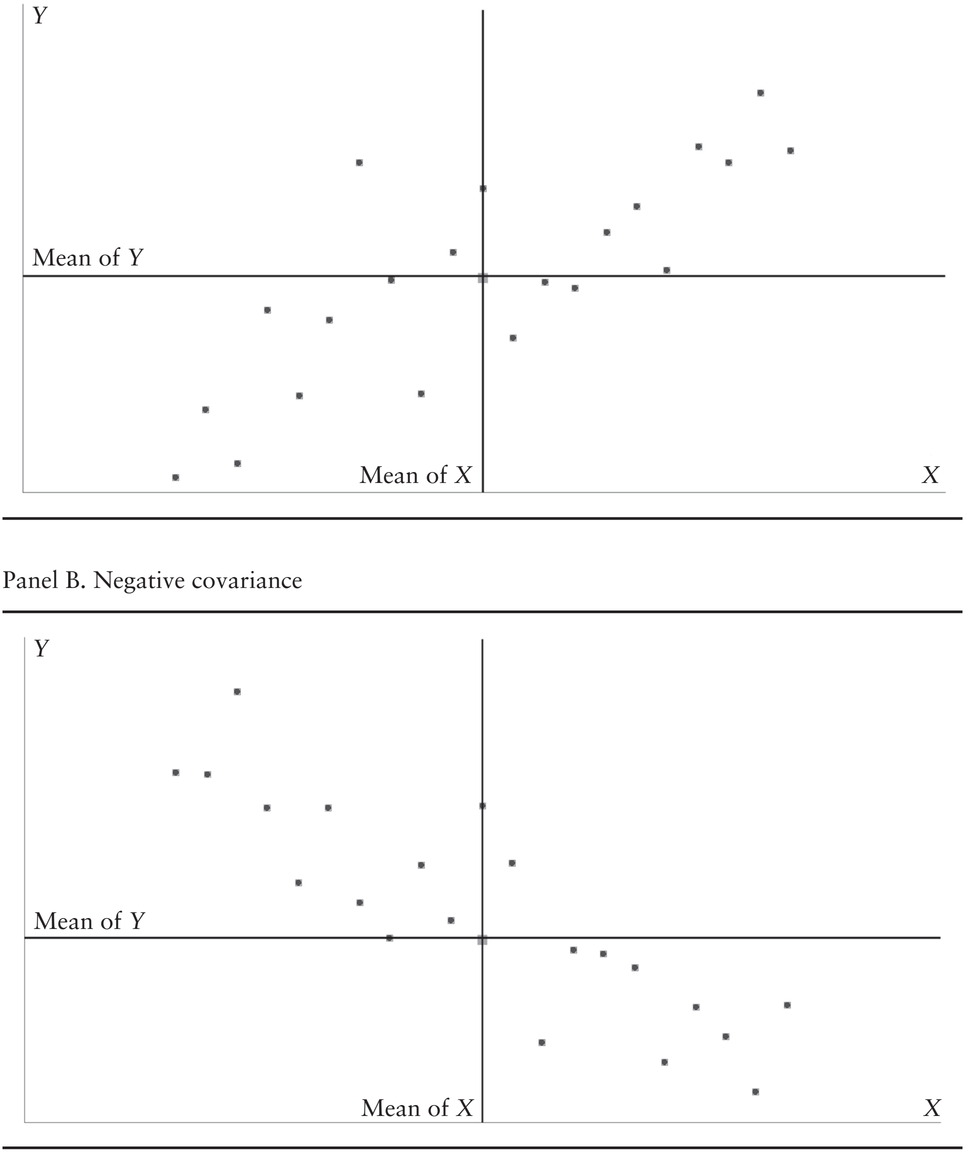
where p_{ij} represents the joint probability of X and Y occurring. The *covariance* is a measure of the linear association between X and Y. Covariance is positive when both variables are above and below their means at the same time and is negative when X is above its mean when Y is below its mean. Figure A.2 shows the association between two variables X and Y when the covariance is positive and negative.

Note that the covariance depends on the units in which X and Y are measured. To make the covariance *scale-free*, the association between X and Y is often expressed in terms of the *correlation coefficient,*

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \tag{A.4}$$

The correlation coefficient always lies between –1 and +1.

FIGURE A.2 Positive and negative covariance between two random variables.
Panel A. Positive covariance



Panel B. Negative covariance

Semivariance and Semi-Standard Deviation

The *semivariance* of a random variable measures the dispersion of the distribution around a constant B_X for only part of the probability distribution. The *lower semivariance*, for example, is

$$\text{Lower semivariance} = E[\min(X - B_X, 0)^2] = \sum_{i=1}^{N} p_i[\min(X_i - B_X, 0)^2] \quad (\text{A.5})$$

The (positive) square root of the semi-variance is called the *semistandard deviation* or, sometimes, the *semideviation*. Lower semistandard deviation of return, where is set equal to the risk-free rate of interest, is a less commonly-used, but more intuitively appealing, risk measure.

Semicovariance and Semicorrelation

Like semivariance is to variance, semicovariance is to covariance. The *lower semicovariance* of X and Y is defined as

$$\begin{aligned}\text{Semicovariance} &= E[\min(X - B_X, 0)\min(Y - B_Y, 0)] \\ &= \sum_{i=1}^{N}\sum_{j=1}^{N} p_{ij}\min(X_i - B_X, 0)\min(Y_j - B_Y, 0)\end{aligned} \quad (\text{A.6})$$

where p_{ij} represents the joint probability of X and Y occurring and B_X and B_Y are the boundaries for variables X and Y. The lower *semicorrelation coefficient* is

$$\text{Lower semicorrelation} = \frac{\text{Lower semicovariance}}{\text{Lower semideviation}_X \text{Lower semideviation}_Y} \quad (\text{A.7})$$

and always lies between –1 and +1.

Skewness

The *skewness* of a random variable measures the degree of asymmetry of the distribution around the mean. The skewness, denoted γ_1, is third standardized moment of the distribution and is defined as

$$\text{Skew}(X) = \gamma_1 = \frac{E[X - E(X)]^3}{\sigma^3} = \frac{1}{\sigma^3}\sum_{i=1}^{N} p_i[X_i - E(X)]^3 \quad (\text{A.8})$$

where σ is the standard deviation of the distribution. Generally speaking, a distribution is positively skewed (right-skewed) if the higher tail is longer and negatively skewed (left-skewed) if the lower tail is longer.

Kurtosis

The *kurtosis* of a random variable measures the degree of the “peakedness” of the distribution around the mean. The kurtosis, denoted γ_2, is fourth standardized moment of the distribution and is defined as

$$Kurt(X) = \gamma_2 = \frac{E[X - E(X)]^4}{\sigma^4} = \frac{1}{\sigma^4}\sum_{i=1}^{N} p_i[X_i - E(X)]^4 \tag{A.9}$$

where σ is the standard deviation. In most statistical software, excess kurtosis rather than kurtosis is reported. *Excess kurtosis* is defined as $\gamma_2 - 3$. For a normal distribution, excess kurtosis equals 0. Positive excess kurtosis implies that the distribution of X is more peaked in the center than the normal and has fatter tails. Such a distribution is said to be “leptokurtic.” Negative excess kurtosis implies that the distribution of X is flatter in the middle and has smaller tails. Such a distribution is said to be “platykurtic.” Finally, when excess kurtosis equals zero (like the normal), the distribution is said to be “mesokurtic.”

PROPERTIES OF EXPECTATION OPERATORS

Many finance applications, particularly those associated with portfolio selection, involve using expectations of the parameters of future security rate of return distributions. Since a security portfolio is nothing more than a weighted sum of its constituent securities, we are interested in understanding how random security returns aggregate into portfolios. Table A.2 presents some key properties of expectations operators. In the table, X and Y are assumed to be random variables, and a and b are assumed to be known constants. In the remainder of this section, we use these results in examining the properties of the formulas we use to estimate the parameters of probability distributions.

ESTIMATION

Means, variances, and covariances are measured with certainty only if we have the population (i.e., all possible outcomes) at our disposal. More typically, however, we have a *sample* from the population and want to make inferences about the population. In this section, assume we have a sample of n data points from

TABLE A.2 Key properties of expectations operators. X and Y are random variables, and a and b are known constants.

$E(aX + b) = aE(X) + b$	(P-1)
$E[(aX)^2] = a^2E(X^2)$	(P-2)
$Var(aX + b) = a^2Var(X)$	(P-3)
$E(X + Y) = E(X) + E(Y)$	(P-4)
$Var(X + Y) = Var(X) + Var(Y) + 2\ Cov(X,Y)$	(P-5)
Also, if X and Y are independent,	
$E(XY) = E(X)E(Y)$	(P-6)
$Cov(X,Y) = 0$	(P-7)

the population. Our objective is to estimate characteristics of the population, and then attempt to draw conclusions about the population parameters. An *estimator* is the formula used to estimate a population parameter; an *estimate* is the value obtained from an estimator for a particular sample.

Estimator of Mean

An estimator is said to be *unbiased* if the expected value of the estimator is equal to the population parameter. The estimator of the sample mean is

$$\hat{\mu}_X = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{A.10}$$

This estimator is unbiased since its expected value equals the population mean, that is,

$$\begin{aligned} E(\hat{\mu}_X) &= E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu_X = \frac{1}{n}n\mu_X = \mu_X \end{aligned}$$

Estimator of Variance

The unbiased estimator of the variance of a random variable is

$$\hat{\sigma}_X^2 = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \hat{\mu}_X)^2 \tag{A.11}$$

The reason $n - 1$ (rather than n) appears in the denominator is that, in order to compute the sample variance, the sample mean must first be computed. This places a constraint on the n data points in the sample. That is, the n observations must sum to n times the computed mean, $\hat{\mu}_X$. This leaves $n - 1$ unconstrained observations with which to estimate the sample variance.

Estimator of Covariance and Correlation

The unbiased estimator for sample covariance is

$$\hat{\sigma}_{XY} = \frac{1}{n-1}\sum_{i=1}^{n} (X_i - \hat{\mu}_X)(Y_i - \hat{\mu}_Y) \tag{A.12}$$

The adjustment to the denominator is made because, in calculating the sum of the products of the deviations in X and Y, there are n observations on the joint outcomes of X and Y and thus n independent pieces of information. One piece of information is used to calculate the means of X and Y, however. The sum of all n observations is constrained to be equal to n times the means of X and Y, respectively. As a result, there are degrees of freedom.

Finally, the *sample correlation coefficient* between the two variables is

$$\hat{\rho}_{XY} = \frac{\sum_{i=1}^{n}(X_i - \hat{\mu}_X)(Y_i - \hat{\mu}_Y)}{\sqrt{\sum_{i=1}^{n}(X_i - \hat{\mu}_X)^2 \sum_{i=1}^{n}(Y_i - \hat{\mu}_Y)^2}} = \frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X \hat{\sigma}_Y} \qquad \text{(A.13)}$$

Estimator of Lower Semivariance

An estimator of sample semivariance is

$$\text{Lower semivariance} = \frac{1}{n}\sum_{i=1}^{n} \min(X_i - B_X, 0)^2 \qquad \text{(A.14)}$$

The (positive) square root of the estimate of the semivariance provides our estimate of the semistandard deviation. In applying (A.12) to return distributions, the most common choices for B_X are the risk-free rate of interest and zero. The choice of the risk-free rate is intuitive in the sense that it says we are only concerned about holding a risky asset to the extent that its return might be below what can be earned by placing the investment funds in a risk-free asset.

Estimators of Semicovariance and Semicorrelation

An estimator of sample semicovariance is

$$\text{Lower semicovariance} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} \min(X_i - B_X, 0)\min(Y_j - B_Y, 0) \qquad \text{(A.15)}$$

where B_X and B_Y are the upper boundaries of variables X and Y. The estimator of the lower semicorrelation coefficient is

$$\text{Lower semicorrelation} = \frac{\text{Lower semicovariance}}{\text{Lower semideviation}_X \text{Lower semideviation}_Y} \qquad \text{(A.16)}$$

and always lies between –1 and +1.

Estimator of Skewness

An estimator of sample skewness is

$$\hat{\gamma}_1 = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{X_i - \hat{\mu}_X}{\hat{\sigma}_X}\right)^3 \tag{A.17}$$

where $\hat{\sigma}_X$ is the estimated standard deviation of the distribution.[2] Positive skewness implies that the distribution has a long tail to the right, and negative skewness implies that it has a long tail to the left. Many financial models incorporate the behavioral assumption that investors gain satisfaction from positive skewness in the rate of return distribution, holding other factors constant.

Estimator of Kurtosis

An estimator of sample excess kurtosis is

$$\hat{\gamma}_2 = \left\{\frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{X_i - \hat{\mu}_X}{\hat{\sigma}_X}\right)^4\right\} - \frac{3(n-1)^2}{(n-2)(n-3)} \tag{A.18}$$

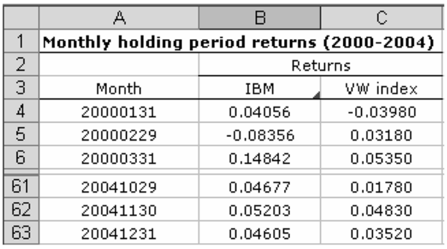
Excess kurtosis characterizes the peakedness or flatness of a distribution relative to the normal distribution. Positive kurtosis indicates a relatively peaked distribution, and negative kurtosis indicates a relatively flat distribution.

ILLUSTRATION A.1 Estimate mean, variance, standard deviation, skewness, and excess kurtosis of monthly stock returns for IBM.

*The worksheet **A1** in the Excel file, **A Illustrations.xls**, contains 60 months of returns for IBM and a value-weighted stock market index over the period January 2000 through December 2004. Estimate the mean, variance, standard deviation, skewness, and kurtosis of IBM's return series. Use the standard Excel statistical functions to perform your computations. Comment on the levels of skewness and kurtosis.*

To begin, examine the contents of the data file illustrated on the following page. The first column contains the date of the month-end. The next two columns contain the rates of return of IBM's stock and a value-weighted stock market index. Note that rows 7 through 60 have been compressed some that the file contents can be displayed on one page. You can adjust the height to see the contents of the cells if necessary.

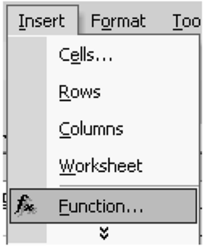
[2] While it is beyond the scope of this appendix, the $1/(n-1)$ allows for the fact that a degree of freedom has been used in estimating the mean, and $n/(n-2)$ is a small sample bias adjustment.



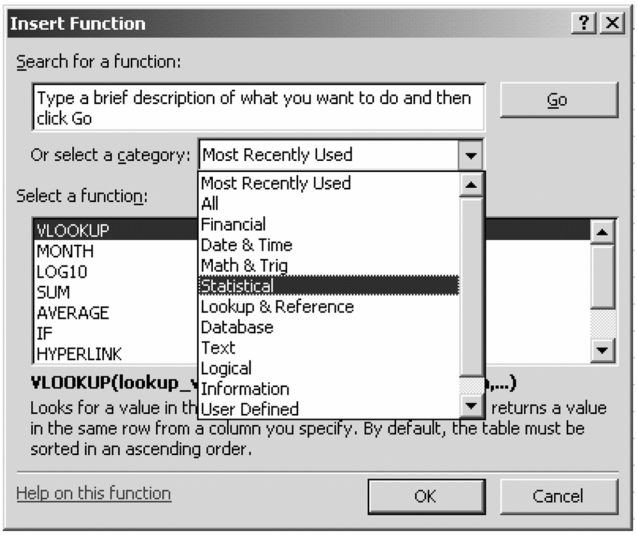
	A	B	C
1	Monthly holding period returns (2000-2004)		
2		Returns	
3	Month	IBM	VW index
4	20000131	0.04056	-0.03980
5	20000229	-0.08356	0.03180
6	20000331	0.14842	0.05350
61	20041029	0.04677	0.01780
62	20041130	0.05203	0.04830
63	20041231	0.04605	0.03520

Rather than perform the computations of each estimator, we will rely on Microsoft Excel add-ins. Some are part of the Excel add-in function library provided by Microsoft. Others are part of the OPTVAL add-in function library that is part of the CD that accompanies this book. The same approach is used to apply the functions from either library.

The first step in applying an add-in function is to click on the "Insert" menu and select "Function" as shown:



Clicking on "Function" will cause a menu to appear. The menu contains the different sub-libraries of add-ins that are available. "All" contains the entire set of add-in functions. It is so lengthy, it is cumbersome to use. In the Insert Function dialog box as shown, "All" are the functions separated into categories according to their general purpose. For this illustration, we need functions from the "Statistical" category.



Insert Function

Search for a function:

Type a brief description of what you want to do and then click Go

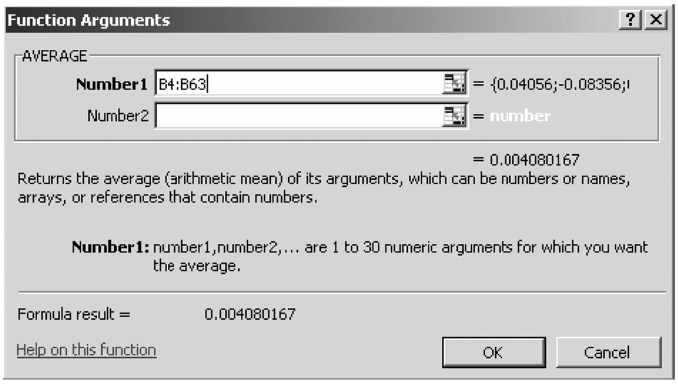
Or select a category: Most Recently Used

Select a function:

VLOOKUP(lookup_v...,...)

Looks for a value in th... returns a value in the same row from a column you specify. By default, the table must be sorted in an ascending order.

Clicking on "Statistical" will provide a list of statistical functions. The "Average" function is used to compute the estimate of the mean using (A-10). When we click on the function name, the following form appears. To insert the IBM return series in computing the mean, simply place the cursor in the Function Arguments dialog box to the right of Number1 and highlight the cells B4 through B63 as shown, and then click “OK”:



Function Arguments

AVERAGE

Number1 B4:B63 = {0.04056;-0.08356;

Number2 = number

= 0.004080167

Returns the average (arithmetic mean) of its arguments, which can be numbers or names, arrays, or references that contain numbers.

Number1: number1,number2,... are 1 to 30 numeric arguments for which you want the average.

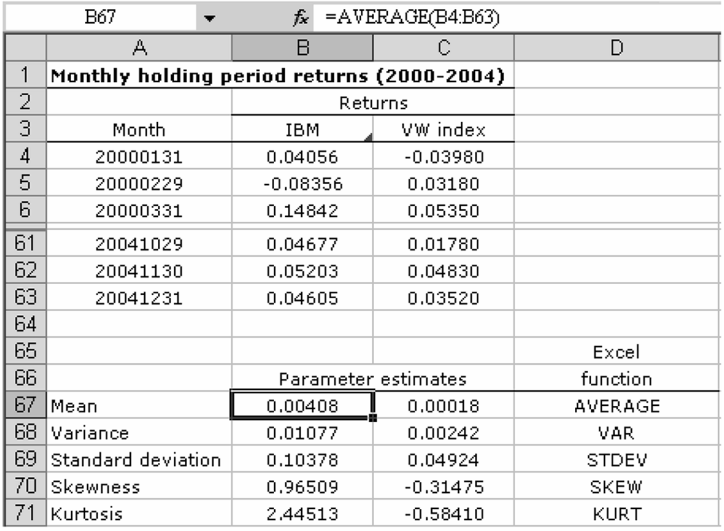
Formula result = 0.004080167

The illustration that follows summarizes the results. Note that the contents of cell B67 is the mean monthly return of IBM, 0.00408. Cell C67 contains the mean return of the market index and involves the function call “=AVERAGE(C4:C63)”.

For your convenience, the Excel function names of all of the remaining estimators are provided in column D. An inspection of the worksheet shows that cells B67 through B71 have the following function calls:

=AVERAGE(B4:B63)
=VAR(B4:B63)
=STDEV(B4:B63)
=SKEW(B4:B63)
=KURT(B4:B63)

The estimated skewness of IBM’s observed monthly returns is 0.96509. Positive skewness implies that the return distribution is asymmetric and has a long tail on the right. The estimated kurtosis is 2.44513. Positive excess kurtosis implies that the return distribution is more peaked than the normal and has fatter tails.



B67 =AVERAGE(B4:B63)

	A	B	C	D
1	**Monthly holding period returns (2000-2004)**			
2		Returns		
3	Month	IBM	VW index	
4	20000131	0.04056	-0.03980	
5	20000229	-0.08356	0.03180	
6	20000331	0.14842	0.05350	
61	20041029	0.04677	0.01780	
62	20041130	0.05203	0.04830	
63	20041231	0.04605	0.03520	
64				
65				Excel
66		Parameter estimates		function
67	Mean	0.00408	0.00018	AVERAGE
68	Variance	0.01077	0.00242	VAR
69	Standard deviation	0.10378	0.04924	STDEV
70	Skewness	0.96509	-0.31475	SKEW
71	Kurtosis	2.44513	-0.58410	KURT

To understand the meaning of the skewness and kurtosis parameter values in relation to the shape of the return distribution, it is useful to plot a *histogram.* A histogram typically divides the distance between the minimum and maximum values of the sample of observations into equal intervals and then tabulates the number of observations that fall within each interval. The lowest monthly return for IBM during the sample period is –22.6% in September 2002 and the highest is 35.4% in October 2002. The total number of monthly returns is 60. In the following figure, we display the frequency distribution of actual monthly returns for IBM during the period (i.e., the light-colored bars). We also shown the frequency of returns that is expected if IBM's returns were normally distributed during the period (i.e. the dark-colored bars).[3] Note that the patterns are just as expected. During the sample period, IBM had more large positive returns and fewer large negative returns relative to a normal distribution. This represents positive skewness. Also, during the sample period, IBM had more instances in which the observed monthly was very close to the mean. The peakedness shown in the histogram represents positive excess kurtosis.

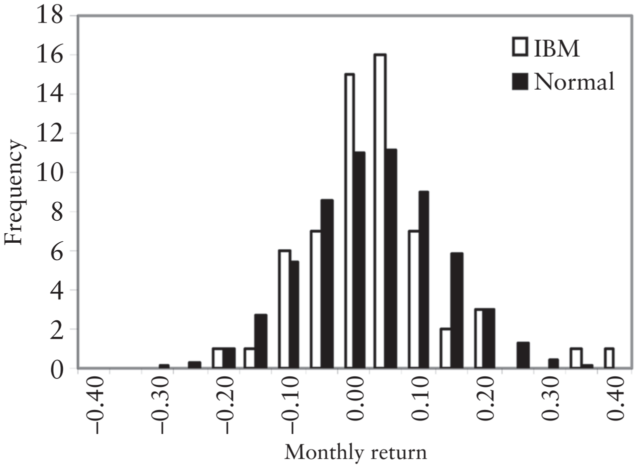
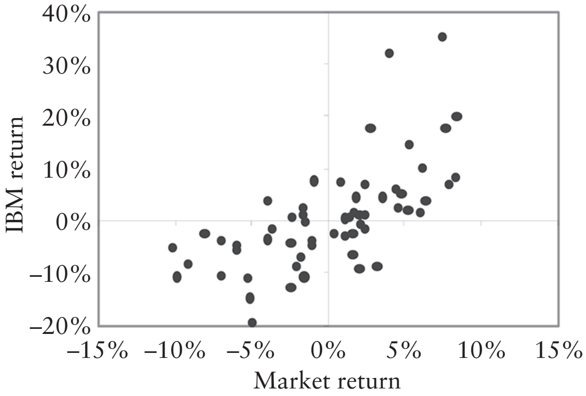


ILLUSTRATION A.2 Estimate covariance and correlation between IBM and stock market index returns.

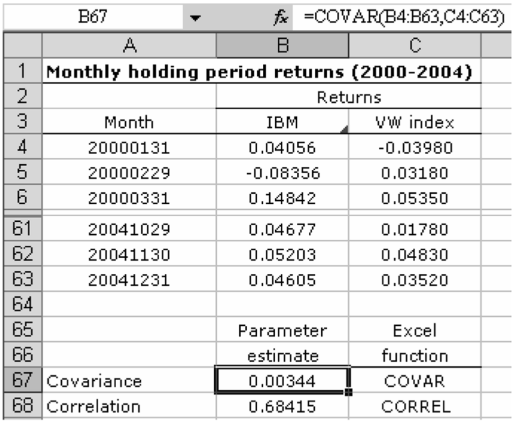
*Using the monthly returns reported in the **A2** worksheet of the Excel file, **A Illustrations.xls**, estimate the covariance and correlation between the IBM and market return series. Use the standard Excel statistical functions to perform your computations.*

As noted earlier in this appendix, covariance and correlation measure the association between two random variables. To get a sense of the relation between two variables, it is useful to plot the series against one another. The figure below shows us that, when the market return is positive, IBM's return is positive, and, when the market return is negative, IBM's return is negative. In other words, the returns of two series are positively correlated (have positive covariance).

[3] The mean and the standard deviation of the normal distribution are set equal to the mean and the standard deviation estimated for the sample, 0.00408 and 0.10378, respectively.



The Excel functions for computing the covariance (A-12) and correlation (A-13) are COVAR and CORREL, respectively. Using the information in the worksheet *A2*, the estimates of covariance and correlation are:



=COVAR(B4:B63,C4:C63)

	A	B	C
1	**Monthly holding period returns (2000-2004)**		
2		Returns	
3	Month	IBM	VW index
4	20000131	0.04056	-0.03980
5	20000229	-0.08356	0.03180
6	20000331	0.14842	0.05350
61	20041029	0.04677	0.01780
62	20041130	0.05203	0.04830
63	20041231	0.04605	0.03520
64			
65		Parameter	Excel
66		estimate	function
67	Covariance	0.00344	COVAR
68	Correlation	0.68415	CORREL

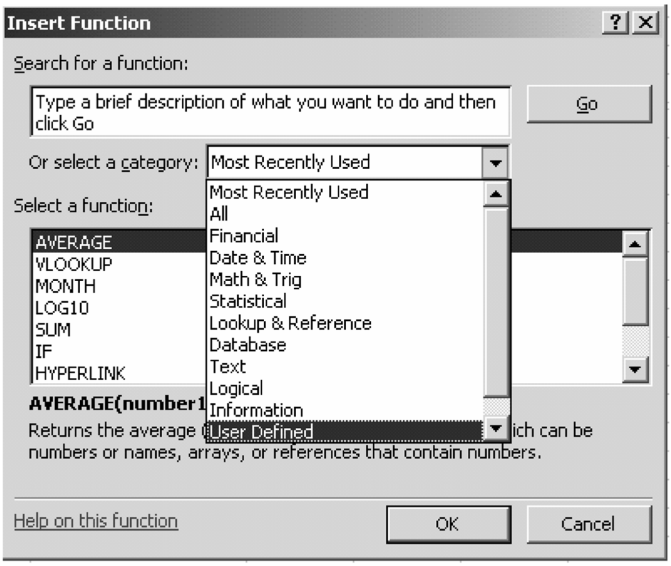
The estimated correlation is 0.68415, which implies that the returns are strongly positively correlated.

ILLUSTRATION A.3 Estimate semivariance and semistandard deviation of return distributions. Also estimate semicovariance and semicorrelation.

*Using the information provided in the worksheet **A3** in the Excel file, **A Illustrations.xls**, estimate semivariance and semistandard deviation of the return series for IBM and the market. Also, estimate semicovariance and semicorrelation between IBM and stock market index returns. Compare the correlation and semicorrelation estimates and comment on the difference.*

The “Statistical” library in Excel contains the most commonly-used statistical functions in applications from all disciplines. This book focuses exclusively on finance applications, and certain useful statistical functions are not included in the Excel statistical library. Consequently, these functions are included in the OPTVAL function library.

To use the OPTVAL functions, we click on the "User Defined" option in the Insert Function menu as shown:



Search for a function: Type a brief description of what you want to do and then click Go

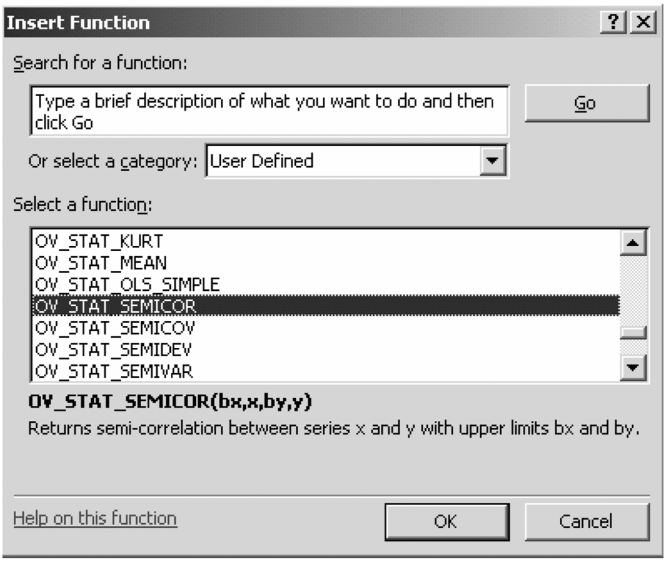
Or select a category: Most Recently Used

Select a function:

AVERAGE(number1

Returns the average (ich can be numbers or names, arrays, or references that contain numbers.

What will appear is the list of user-defined functions. They are clustered together in the menu by virtue of the fact that they begin with the prefix "OV_". The next term in the function name describes the category. The statistical functions in the OPTVAL library begin with "OV_STAT_" as shown:



Search for a function: Type a brief description of what you want to do and then click Go

Or select a category: User Defined

Select a function: OV_STAT_SEMICOR

OV_STAT_SEMICOR(bx,x,by,y)

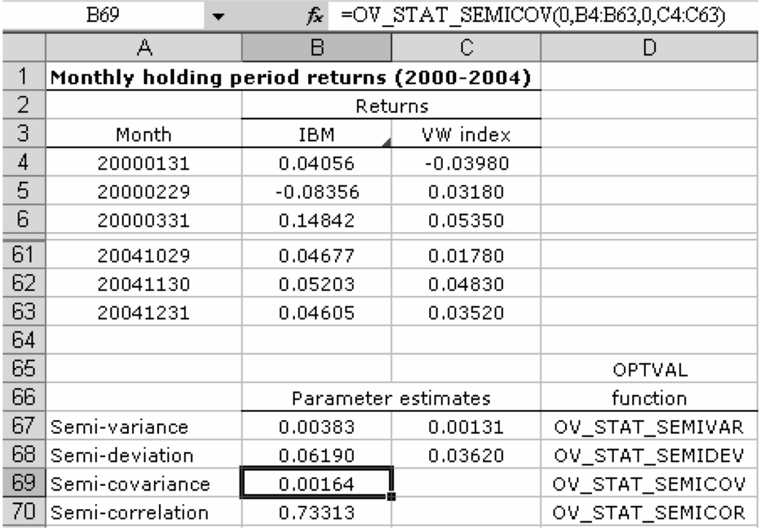
Returns semi-correlation between series x and y with upper limits bx and by.

The remaining part of the name corresponds to the nature of the computation. The worksheet below illustrates the use of the lower semi-correlation function. The syntax of the function is

$$\text{OV\_STAT\_SEMICOR}(bx, x, by, y)$$

where *bx* is the upper bound on the observations of *x*, *x* is a vector containing the observations of *x*, *by* is the upper bound on the observations of *y*, and *y* is the vector of *y* observations.

A summary of the computations is contained in the illustration that follows. Interestingly, the lower semicorrelation estimate, 0.73313, is greater than the correlation, 0.68415. The fact that both correlations are positive indicates that IBM and the market tend to move together, however, the fact that the lower semicorrelation is higher means that the relation is strongest when prices fall. This is type of behavior is not uncommon in financial markets. A declining market sometimes causes investors to leave a particular asset class in favor of a safer one (e.g., sells stocks and buy Treasury bills).



B69 fx =OV_STAT_SEMICOV(0,B4:B63,0,C4:C63)

	A	B	C	D
1	**Monthly holding period returns (2000-2004)**			
2		Returns		
3	Month	IBM	VW index	
4	20000131	0.04056	-0.03980	
5	20000229	-0.08356	0.03180	
6	20000331	0.14842	0.05350	
61	20041029	0.04677	0.01780	
62	20041130	0.05203	0.04830	
63	20041231	0.04605	0.03520	
64				
65				OPTVAL
66		Parameter estimates		function
67	Semi-variance	0.00383	0.00131	OV_STAT_SEMIVAR
68	Semi-deviation	0.06190	0.03620	OV_STAT_SEMIDEV
69	Semi-covariance	0.00164		OV_STAT_SEMICOV
70	Semi-correlation	0.73313		OV_STAT_SEMICOR

PROBABILITY DISTRIBUTIONS

In the remainder of this appendix, we work with four specific *continuous* density functions—the normal, chi-squared, *t*, and *F* distributions.[4] Unlike a discrete density function, a continuous random variable can take on any value from the real number line from $-\infty$ to $+\infty$. We use the normal distribution to develop measures of risk. We use the remaining three distributions to help develop a framework for understanding the role of measurement error in security valuation and risk measurement.

Normal Distribution

The *normal distribution* is important for a number of reasons. First, it is symmetric and bell-shaped, and closely approximates many empirical distributions such as security returns and cash flows. Second, it is fully described by its mean

[4] In Chapter 7, we also use the log-normal distribution in describing the distribution of future security prices.

and variance, so we need not worry about other properties such as skewness and kurtosis. Third, if two (or more) random variables are normally distributed with identical means and variances, any weighted sum of these variables will be normally distributed.

The *normal distribution* is a continuous bell-shaped probability distribution whose density function is given by

$$p(X_i) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{\left[-\frac{1}{2\sigma_X^2}(X_i - \mu_X)^2\right]} \tag{A.19}$$

where μ_X and σ_X are the mean and standard deviation of X. In the special case where $\mu_X = 0$ and $\sigma_X = 1$, the resulting random variable (usually denoted z) has a *standard normal density function*,

$$n(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \tag{A.20}$$

Figure A.3, Panel A plots $n(z)$ as a function of z. Note that all normal distributions can be be transformed into the standard (or unit) normal distribution using the relation, $z_i = (X_i - \mu_X)/\sigma_X$.

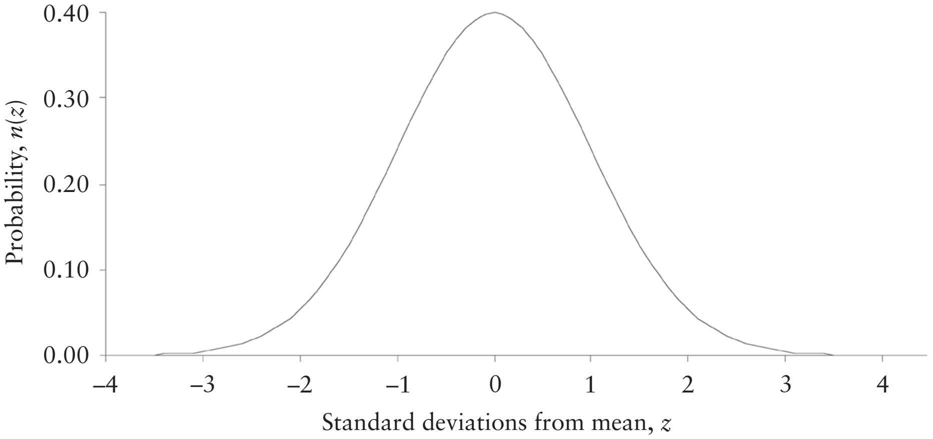
To compute the probability that a random drawing from a standard normal distribution will fall below a level a, we integrate (A.20) over the range from $-\infty$ to a, that is,

$$\begin{aligned} \Pr(\tilde{z} < a) &= \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= N(a) \end{aligned} \tag{A.21}$$

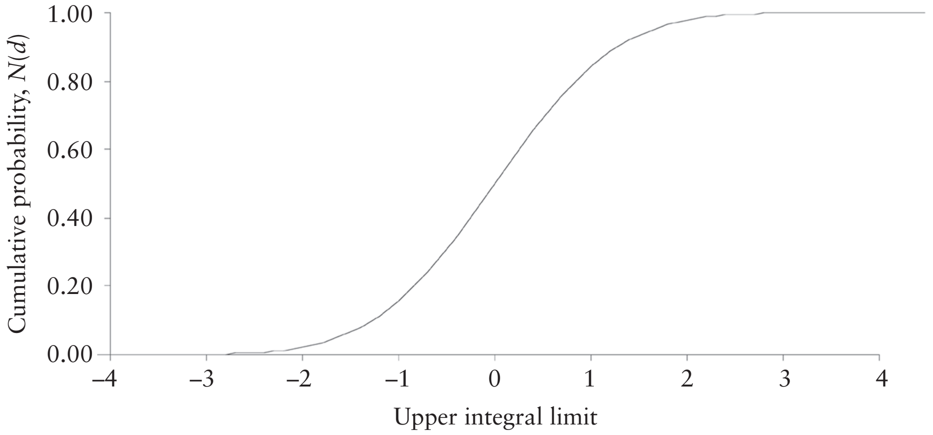
The usual way in which values of $N(a)$ have been available in matrices like Tables C.1A and 1B in Appendix C of this book. Appendix C contains all of the statistical tables that we will need in hypothesis testing and building confidence intervals. In Table C.1A, for example, $N(-2.00) = 0.0228$. This means that the chance that a random drawing from a standard normal distribution will have a value more than two standard deviation below the mean is 2.28%. Since the standard normal distribution is symmetric and centered on 0, this also means that the chance that a random drawing from a standard normal distribution will have a value more than two standard deviation above the mean is 2.28%. To check this, we can turn to Table C.1B, where we find that $N(2.00) = 0.9772$, that is, the chance that a random drawing from a standard normal distribution will have a value less than two standard deviations above the mean is 97.72%. The complement of this value is, of course, 2.28%. The chance that a random drawing from a standard normal distribution will have a value in the range plus or minus two standard deviations from the mean is 97.72 – 2.28 = 95.44%.

FIGURE A.3 Standard normal distribution function and cumulative standard normal density function.

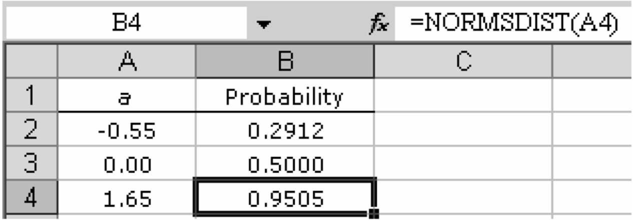
Panel A. Standard normal



Panel B. Cumulative standard normal



In recent years, commonly used statistical software packages have begun to include functions for evaluating the integral (A.21). Microsoft Excel, for example, has an add-in function called NORMSDIST that computes the cumulative standard normal probability, $N(a)$. The following illustration shows how the function is called as well as sample values. Note that the values correspond to the values reported in Tables C.1A and 1B. Figure C.1, Panel B shows the cumulative probability $N(a)$ as a function of a.



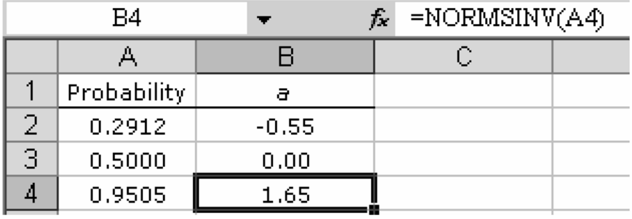
B4 fx =NORMSDIST(A4)

	A	B	C
1	a	Probability	
2	-0.55	0.2912	
3	0.00	0.5000	
4	1.65	0.9505	

Closely related to the NORMSDIST function is the NORMSINV function, which computes the inverse of the cumulative standard normal density function. Suppose we are interested in determining the level of *a* that makes the cumulative probability equal to 5%, that is,

$$\int_{-\infty}^{a^*} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = 0.05$$

Using *a* = 0.05 in the inverse function shows that NORMSINV(0.05) = –1.645. This imples that the chance of a random drawing from a standard normal distribution producing a value at or below –1.645 standard deviations below the mean is 5%. Alternatively, it implies that we are 95% confident that a random drawing from a standard normal distribution will produce a value exceeding –1.645. This illustration shows sample functions calls and values:



B4 fx =NORMSINV(A4)

	A	B	C
1	Probability	a	
2	0.2912	-0.55	
3	0.5000	0.00	
4	0.9505	1.65	

ILLUSTRATION A.4 Compute maximum possible loss over next month with 95% confidence.

Assume you hold $10 million of IBM's stock as of December 31, 2004. Based on the returns that appear in the worksheet ***A4*** *in the Excel file,* ***A Illustrations.xls,*** *compute the expected maximum (or "worst loss") that we can expect to occur over the next month with 95% confidence. How does the result change if you assume IBM's returns are normally distributed?*

As a risk manager, you will be often placed in situations in which you will need to quantify the level of risk you face. There are a variety of ways to go about this task, and we will discuss several in the chapters of the text. The one discussed here is called *Value-at-Risk* or simply *VAR*. What VAR attempts to measure is the maximum dollar loss we can expect to incur over the given period of time at a particular confidence level.

Empirical Distribution

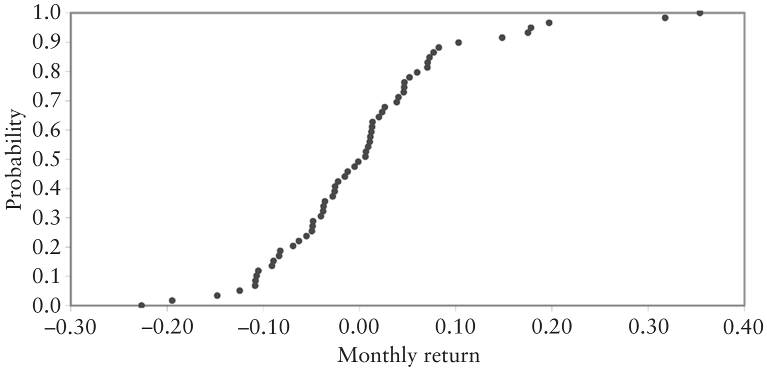
One way we can go about estimating this quantity is to use the realized *empirical distribution*, that is, the distribution of returns as they appeared in the recent past. The intuition is that, unless there is reason to believe otherwise, the next observed return should be drawn from the same distribution.

The worksheet ***A4*** contains the most recent 60 months of IBM stock returns. Each return in the series is assumed to have an equal chance of occurring again. Suppose we order the returns from lowest to highest. With 60 return observations, the number of intervals between observations is 59. Hence, the probability of falling into a particular interval is 1/59 or 1.695%. The first few observations in the ordered return series together with their receptive probabilities are:

Monthly holding period returns (2000–2004)

Month	IBM return	Cumulative probability
20020930	–0.22645	0
20020430	–0.19462	0.01695
20000929	–0.14773	0.03390
20001031	–0.12444	0.05085
20021231	–0.10838	0.06780
20020131	–0.10805	0.08475

Since no return below –22.645% appeared in the 60-month history, the probability that a drawing from this distribution will have a value below –22.645% is 0.[5] The probability that a drawing from this distribution will have a value below –12.444% is 0.05085. The cumulative probability function for this empirical distribution is:



The question is, however, what is the critical return below which there is a 5% chance of occurrence. Looking at the above table, the critical return lies somewhere in the range between –14.773% and –12.444%. To find exactly where, we interpolate using the cumulative probabilities as weights, that is,

$$-14.773\left(\frac{0.05085-0.05000}{0.05085-0.03390}\right)-12.444\left(\frac{0.05000-0.03390}{0.05085-0.03390}\right)=-12.560\%$$

In other words, based on the empirical distribution of IBM's returns, the chance of experiencing a return of –12.560% or less over the next month is 5%. Alternatively, we are

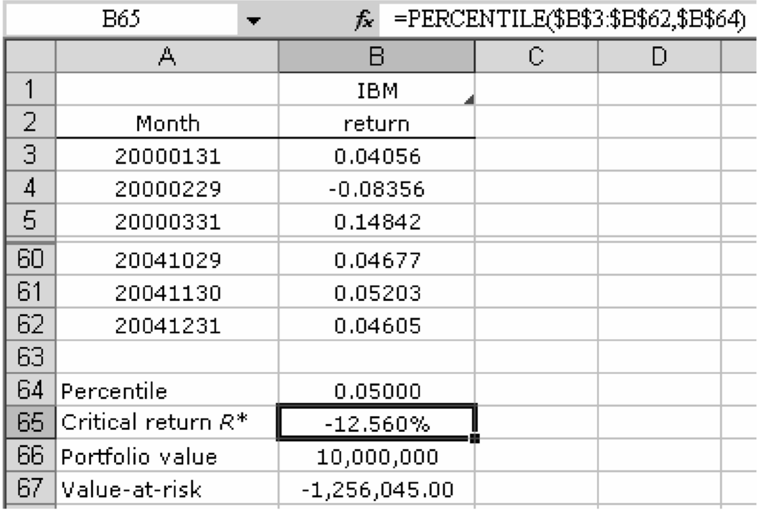
[5] The fact that no return below –22.645% has been observed does not mean that no returns will ever fall below that level. This is a weakness of using the empirical distribution approach to estimating VAR.

95% confident that the worst loss we will experience over the next month is –12.560% of the portfolio value or $1,256,045.

As it turns out, another Excel statistical function can compute this critical return directly. The syntax of the function is

PERCENTILE(*array*,*k*)

where *array* is the vector of monthly returns and *k* is the probability level. In using this function, there is no need to arrange the monthly return series in ascending order. In the event that the critical return falls between observed returns (as it does in this illustration), the function performs the interpolation automatically. To verify this result, consider the following:



B65 fx =PERCENTILE($B$3:$B$62,$B$64)

	A	B	C	D
1		IBM		
2	Month	return		
3	20000131	0.04056		
4	20000229	-0.08356		
5	20000331	0.14842		
60	20041029	0.04677		
61	20041130	0.05203		
62	20041231	0.04605		
63				
64	Percentile	0.05000		
65	Critical return *R**	-12.560%		
66	Portfolio value	10,000,000		
67	Value-at-risk	-1,256,045.00		

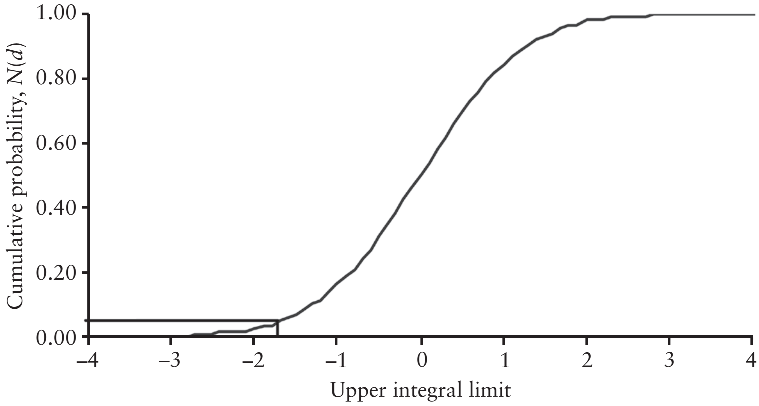
Normal distribution

A second approach to estimating value-at-risk is to assume that security returns have a parametric distribution. The most common assumption in this regard is that returns are normal distributed. Consequently, the only parameters we need to characterize the distribution are the mean and the standard deviation. To find these values, we rely the historical returns, and then work with the mechanics of the normal distribution to do the rest.

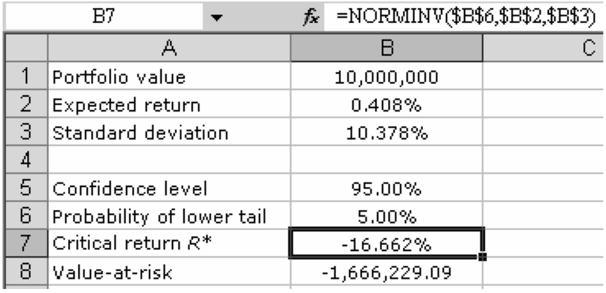
The mean and standard deviation of IBM returns over the sample period were 0.00408 and 0.10378, respectively. From the discussion of the standard normal distribution earlier, we know that we can use the NORMSINV function to find the critical value of a^* such that $n(a^*) = 0.05$, as shown in the following figure. From an earlier illustration, we know that the critical value of a^* is –1.645. Thus, the critical return (i.e., the worst loss over the next month with 95% confidence) is

$$R^* = 0.00408 - 1.65(0.10378) = -0.16662$$

and the VAR under the assumption of normally distributed returns is $1,666,200. This number exceeds the VAR under the empirical distribution because the empirical distribution is positively skewed. The normal distribution assigns a greater chance of large negative returns.

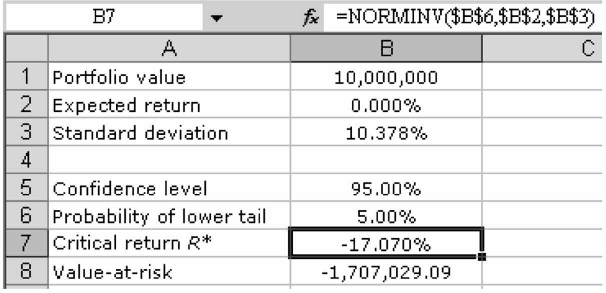


It is also worth nothing that we need not compute the critical return R^* by hand, as we did above. Excel has add-in functions, NORMDIST and NORMINV, that allow the user to prespecify the mean and standard deviation of the normal distribution directly. Thus where NORMSINV returns the critical value of a^* where the mean and standard deviation are 0 and 1, respectively, NORMINV returns the critical value of R^* where the mean and standard deviation are $\hat{\mu}_R$ and $\hat{\sigma}_R$, respectively. Applying the problem parameters, we get:


B7 fx =NORMINV($B$6,$B$2,$B$3)

	A	B	C
1	Portfolio value	10,000,000	
2	Expected return	0.408%	
3	Standard deviation	10.378%	
4			
5	Confidence level	95.00%	
6	Probability of lower tail	5.00%	
7	Critical return *R**	-16.662%	
8	Value-at-risk	-1,666,229.09	

Finally, it is worth noting that VAR is generally defined as the dollar loss relative to the mean. In some instances, however, users prefer to define VAR as the *absolute* dollar loss relative to 0, with no reference to expected value. We can easily accommodate this convention by setting the mean equal to 0 in the above spreadsheet. The absolute dollar VAR is about $1.7 million.


B7 fx =NORMINV($B$6,$B$2,$B$3)

	A	B	C
1	Portfolio value	10,000,000	
2	Expected return	0.000%	
3	Standard deviation	10.378%	
4			
5	Confidence level	95.00%	
6	Probability of lower tail	5.00%	
7	Critical return *R**	-17.070%	
8	Value-at-risk	-1,707,029.09	

Chi-Square Distribution

The *chi-square distribution* plays a key role in many statistical tests. One important application is in the context of answering the question: "Are two sets of data drawn from the same distribution function?" In Illustration A.4, for example, can we test whether the sample of IBM stock returns are drawn from a normal distribution? Below we define the chi-square distribution and its probabilities, and then apply it in tests for distributional differences.

Formally defined, a variable that is the sum of the squares of *n* independent drawings from a standard normal distribution, that is,

$$\chi_n^2 = X_1^2 + X_2^2 + \cdots + X_n^2 \tag{A.22}$$

is said to have a the chi-square distribution with *n* degrees of freedom. The shape of the distribution changes with the number of degrees of freedom, as is shown in Figure A.4. With few degrees of freedom, the distribution is highly positively skewed. As the number of degrees of freedom grows large, the distribution becomes more and more symmetric.

Table A.4 reports the probability that the sum of squared of *n* random standard normal variables will be greater than the critical value χ_n^2. To interpret the table, consider the case where the number of degrees of freedom is 10 and the probability level α is 0.05. The critical χ^2 value is 18.31. This means that the chance of observing a sample χ_{10}^2 value exceeding 18.31 is less than 5%, or, alternatively, we are 95% confident that the sample χ_{10}^2 will be less than 18.31. Figure A.5 illustrates. The darkened tail to the right contains 5% of the area under the χ^2 distribution. The lower bound of this tail is the critical value 18.31. It is also worth noting that Excel has a statistical function that computes the critical value of χ_n^2. Its syntax is

FIGURE A.4 Chi-square (χ^2) distribution with various degrees of freedom.

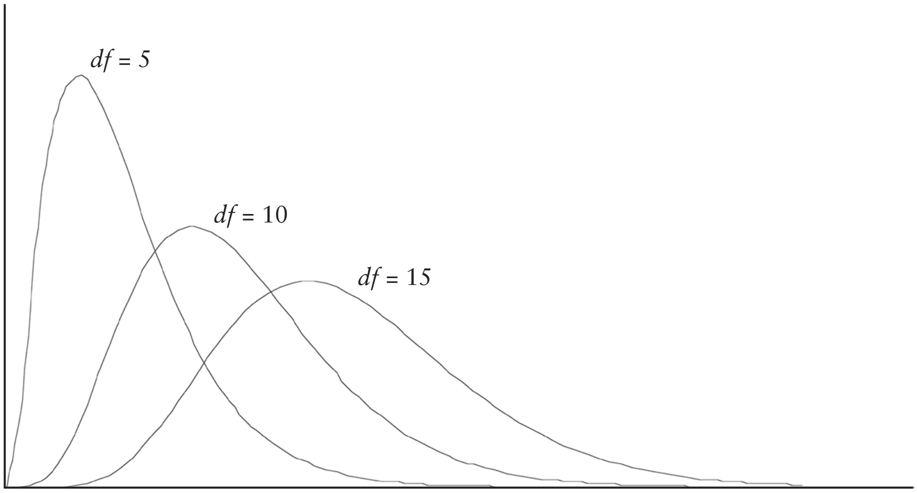
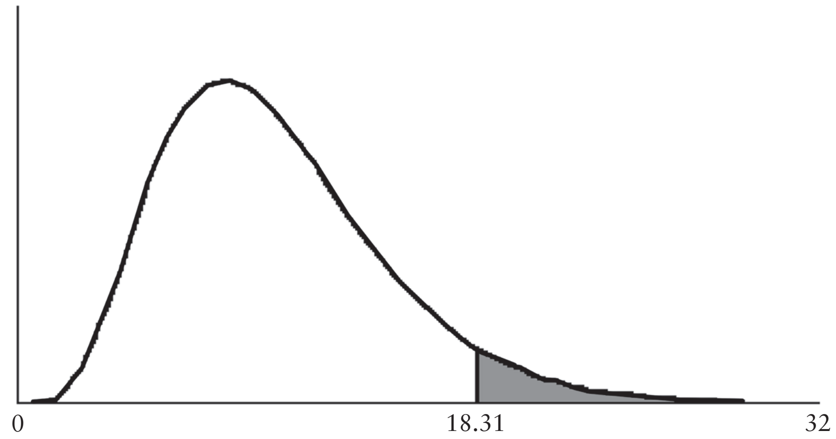


FIGURE A.5 Critical chi-square value at 5% probability level and 10 degrees of freedom.



CHINV(*Probability, Deg_freedom*)

where *Probability* is the chosen significance level, and *Deg_freedom* is the number of degrees of freedom. CHINV(0.05,10) = 18.31.

Tests for Normality One particularly important application of the chi-square distribution is in tests of normality. One simple way of distinguishing between the distributions of two samples is to compute the statistic,

$$\chi^2 = \sum_{i=1}^{n} \frac{(F_i - f_i)^2}{f_i} \tag{A.23}$$

where n is the number of bins, F_i is the number of events observed in the ith bin, and f_i is the expected number under some known distribution such as the normal.[6] In this particular case, the terms in (A.23) are not individually normal, however, if either the number of bins is large or the number of events in each bin is large, the chi-square probability function is a good approximation to the distribution of (A.23). To test the null hypothesis that the sampling distribution is normal, we compute the test statistic (A.23) and compare the value against the critical values reported in Table C.2 in Appendix C.

This is the first of many hypothesis tests that we will perform in this appendix. It is important to note that, *before* any testing is done, we must preset the desired level of significance of our test. The choice of the level of significance, denoted by α, represents the probability of rejecting the null hypothesis when the null hypothesis is, in fact, true. It is our choice, however, conventional levels in statistical analyses are 5% or 1%.

[6] As a practical matter, any term in (A.19) where $n_i = 0$ is ignored.

ILLUSTRATION A.5 Test for normality of stock market returns.

The worksheet A5 in the Excel file, ***A Illustrations.xls****, contains 60 months of returns for a value-weighted stock market index over the period January 2000 through December 2004. Test the null hypothesis that these returns were drawn from a normal distribution.*

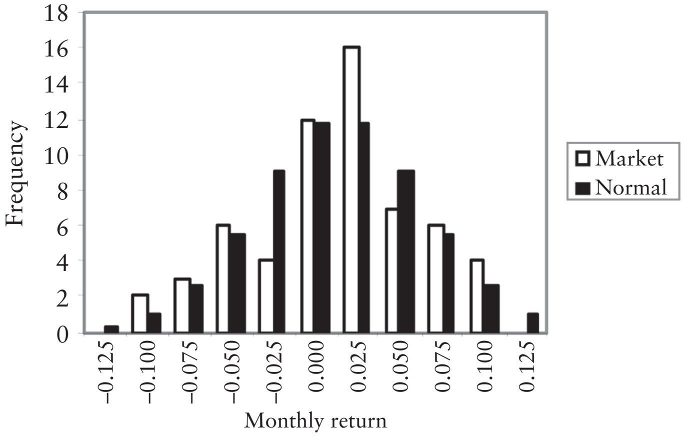
The first step in performing such a test using binned data is to create the binned data. In creating binned data, it is useful to begin with an understanding of the distributions summary statistics. For the 60-monthly market index returns:

	Parameter Estimate	**Excel Function**
Mean	0.00018	AVERAGE
Standard deviation	0.04924	STDEV
Minimum	–0.10250	MIN
Maximum	0.08390	MAX

The range of monthly returns is from –10.250% to 8.390%.

The choice of bins is arbitrary. Based upon the range of observations, we will define the bins to be in 2.5% increments and the range to be from –12.5% to 12.5%. With the bins defined, we then count the number of observations in each bin, that is, identify the F_i's, $i = 1, \ldots, 11$ for use in (A.23).

Next we need to identify the number of observations expected in each bin assuming the monthly returns are normally distributed. The first bin includes all monthly return observations below –12.5%. Under a normal distribution with mean 0.018% and standard deviation 4.924%, the probability of drawing a return below –12.5% is 0.0055. With 60 total return observations, the expected number to fall in this first category is $f_1 =$ 0.330. Note that this value need not be integer. The second bin includes all monthly return observations between –12.5% and –10.0%. Under a normal distribution with mean .018% and standard deviation 4.924%, the probability of drawing a return between –12.5% and –10.0% is 0.0154. With 60 total return observations, the expected number to fall in this second category is $f_2 = 0.926$. The remaining cells in the column are computed in the same manner. The frequencies of observed versus expected numbers of observations in each bin is as follows:



Finally, we compute the individual terms in (A.23) and sum. The computed chi-square value is 8.385. Comparing this value to the critical values reported for 11 degrees of freedom in Table C.2, we find that it lies somewhere between the 10 and 90 percentile values. In other words, we cannot reject the hypothesis that the value-weighted market returns were drawn from a normal distribution. Excel also has a function for computing the chi-square probability. Its syntax is

CHIDIST(*x, deg_freedom*)

where *x* is the computed chi-square value and *deg_freedom* is the number of degrees of freedom. In the current illustration, CHIDIST(8.385,11) = 0.6784.

Before proceeding further, it is important to digress and discuss the concept of a *p*-value, which we have just applied in Illustration A.5 (i.e., the CHIDIST function computes the *p*-value for a χ^2 distribution). As we have noted, the standard procedure for reporting the statistical significance of results of hypothesis testing is to compare the test statistic to the critical value determined at the 5% or 1% significance level. In recent years, however, it has become more common to report *p*-values (probability values). A *p*-value describes the exact significance level associated with a particular test statistic. Thus, a *p*-value of 0.6784 indicates that a coefficient is statistically significant at the 0.6784 level. In the context of a chi-square test with 11 degrees of freedom, this means that 67.84% of the χ^2 distribution lies above 8.385. For purposes of hypothesis testing, we compare the *p*-value with our demanded level of significance, say, $\alpha = 0.05$. Since $0.6784 > 0.05$, we cannot reject the null hypothesis that the market return distribution is normal. Rejection requires that the *p*-value is less than α.

The test statistic (A.23) is useful in demonstrating the intuition underlying why a chi-square test is useful in distinguishing whether there are meaningful differences between the underlying distributions of two samples of data. In the practice, however, we frequently have data that are drawn from continuous distributions. Arbitrarily grouping data into bins involves loss of information. In addition, the selection of bins is arbitrary. For this reason, a considerable amount of energy has been devoted to develop alternative statistics for testing whether a particular sample is drawn from a normal distribution. One well-known test for normality is the Jarque-Bera (1980, 1987) statistic:

$$JB = \frac{n}{6}[\hat{\gamma}_1^2 + \hat{\gamma}_2^2/4] \tag{A.24}$$

where n is the number of sample observations and $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are the sample skewness (A.17) and excess kurtosis (A.18), respectively. The *JB* statistic follows a chi-square distribution with 2 degrees of freedom. If the *JB* statistic is greater than the critical value of the chi-square, we reject the null hypothesis of normality.

ILLUSTRATION A.6 Jarque-Bera test for normality of stock market returns.

The worksheet ***A6*** *of the Excel file,* ***A Illustrations.xls,*** *contains 60 months of returns for a value-weighted stock market index over the period January 2000 through December 2004. Test the null hypothesis that the returns were drawn from a normal distribution using the Jarque-Bera test statisitic.*

To compute the Jarque-Bera test statistic, we need estimates of the skewness and excess kurtosis of the return distribution. Using the appropriate Excel functions for computing (A.17) and (A.18), we find $\hat{\gamma}_1 = -0.31475$ and $\hat{\gamma}_2 = -0.58410$. Thus, the JB statistic is

$$JB = \frac{60}{6}(-0.31475^2 + (-0.58410)^2/4) = 1.8436$$

At 2 degrees of freedom, the sample χ^2 lies in the range between the 10 and 90 percentiles, which means we cannot reject the null hypothesis that the market returns are normally distributed. This conclusion can be confirmed using the Excel function, CHIDIST(1.8436,2) = 0.3978.

t-Distribution

The *Student t-distribution*[7] or, simply, *t-distribution* also plays a key role in statistical analyses. We know from the discussion thus far in this appendix that, in general, we are interested in knowing the parameters of a population but we can neither (a) observe the parameters directly nor (b) observe all of the elements in the distribution. Consequently, we rely upon a sample of observations and statistical analysis to infer the population parameters. The sample mean (A-10), for example, is our "best guess" of the population mean, however, it is a guess. The *t-distribution* helps us quantify the accuracy with which the sample mean estimates the population (or "true") mean.

The random variable,

$$t = \frac{z}{\sqrt{Z/N}} \tag{A.25}$$

is said to have a t-distribution with N degrees of freedom if (a) z is normally distributed with mean 0 and variance 1, (b) Z is distributed as chi-square with N degrees of freedom, and (c) X and Z are independent. Like the standard normal distribution, the t-distribution is symmetric. Unlike the normal distribution, the t-distribution has fat tails when the number of degrees of freedom is small. Figure A.6 illustrates. Although both are centered at 0, the t-distribution has greater variance.

Table C.3 in Appendix C contains percentiles of the t-distribution. The panel heading, *Probability*, is probability that a positive t value will exceed each number in the table in absolute value and is therefore appropriate in one-tailed test. See Figure A.7, Panel A. For a one-tailed test with 10 degrees of freedom and a significance level of $\alpha = 0.05$, the critical t-value t_α is 1.812, that is, the probability that the t-value exceeds 1.812 in absolute value is 5%. For a two-tailed test with 10 degrees of freedom and a significance level of $\alpha = 0.05$, the critical t-value $t_{\alpha/2} = 2.228$, that is, the probability that the t-value is below –2.228 *or* above 2.228 is 5%—2.5% in each tail. See Figure A.7, Panel B.

[7] The t-distribution was derived by William Sealey Gosset in 1908 while he was working at the Guinness brewery in Dublin. He was not allowed to publish under his own name, so the paper was written under the pseudonym "Student." See Student (1908) and http://en.wikipedia.org/wiki/Student%27s_t-distribution.

FIGURE A.6 Student *t*-distribution versus normal distribution.

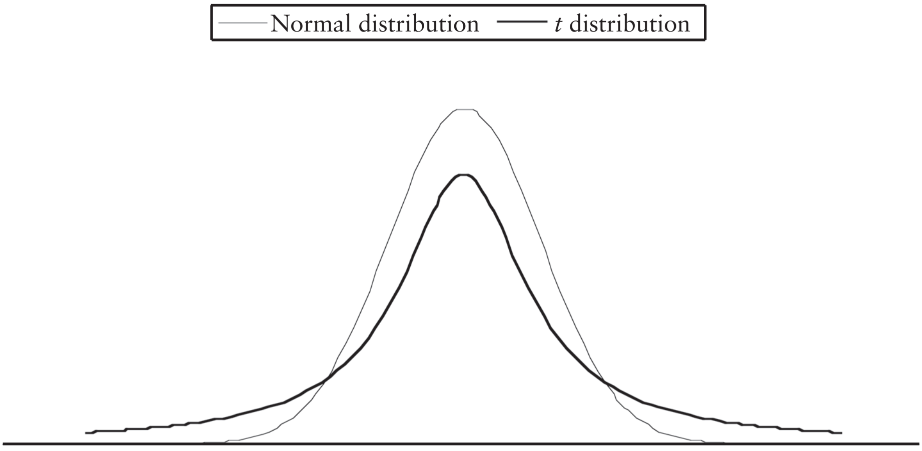
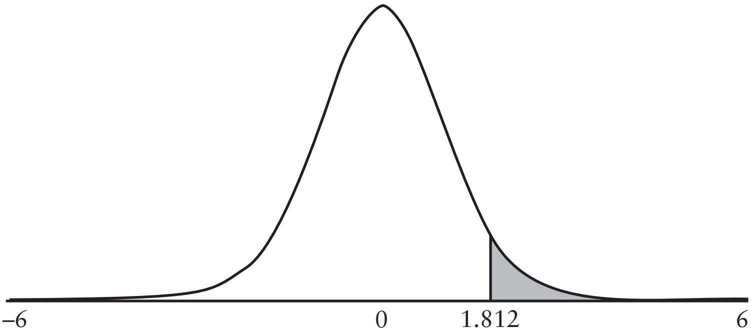
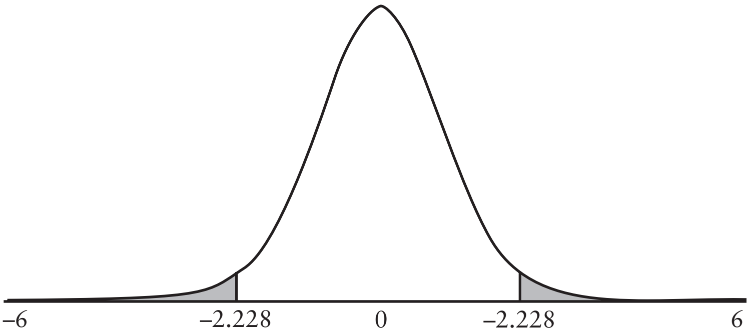


FIGURE A.7 Critical values *t*-distribution at 10 degrees of freedom for one-tailed and two-tailed tests at the 5% level.
Panel A. One-tailed test.



Panel B. Two-tailed test.



To understand how (A.21) helps us, note that the variance of the sample mean is

$$\begin{aligned} Var(\hat{\mu}_X) &= Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) \\ &= \frac{1}{n^2} n\sigma_X^2 = \frac{1}{n}\sigma_X \end{aligned} \tag{A.26}$$

where n is the sample size. The standard deviation of the sample mean is therefore

$$\sigma_{\hat{\mu}_X} = \sigma_X / \sqrt{n} \tag{A.27}$$

Recall that any linear combination of normal distributions is a normal distribution. If X is normally distributed with mean μ_X and standard deviation σ_X, then

$$\frac{\hat{\mu}_X - \mu_X}{\sigma_{\hat{\mu}_X}} = \frac{\hat{\mu}_X - \mu_X}{\sigma_X / \sqrt{n}} \tag{A.28}$$

is normally distributed with mean 0 and standard deviation 1. We use (A.28) in the numerator of (A.25).

Focusing now on the denominator of (A.25), we know that $(n-1)\hat{\sigma}_X^2 / \sigma_X^2$ follows a chi-square distribution with $n - 1$ degrees of freedom. Combining results in (A.25) and simplifying, we find that

$$t = \frac{\dfrac{\hat{\mu}_X - \mu_X}{\hat{\sigma}_X / \sqrt{n}}}{\sqrt{\dfrac{(n-1)\hat{\sigma}_X^2}{\sigma_X^2} / (n-1)}} = \frac{\hat{\mu}_X - \mu_X}{\hat{\sigma}_X / \sqrt{n}} = \frac{\hat{\mu}_X - \mu_X}{s_{\hat{\mu}_X}} \tag{A.29}$$

has a t-distribution with n degrees of freedom. Consequently, we can test whether the mean of a random variable is equal to any particular number using the rightmost term in (A.29), even when the variance of the random variable is unknown. The denominator in the expression, $s_{\hat{\mu}_X}$, is called the *standard error of the estimate*. Note that the standard error becomes small as the sample size grows large. The intuition for this result is that, the more information you gather in estimating the mean, the more reliable your estimate will be.

Test for Zero Mean Perhaps the most common use of the t-statistic is in testing the null hypothesis that the mean of the population is different from zero. Such a test is a special case of (A.29), that is,

$$t = \frac{\hat{\mu}_X}{s_{\hat{\mu}_X}} \tag{A.30}$$

The $(1 - \alpha)\%$ *confidence interval* for the mean of the population is

$$\mu_X \leq \hat{\mu}_X + t_{\alpha,\, df} s_{\hat{\mu}_X} \tag{A.31}$$

where $t_{\alpha,df}$ is the critical t-value corresponding to df degrees of freedom and a desired level of probability α (or desired level of confidence, $1 - \alpha$).

ILLUSTRATION A.7 Test hypothesis mean is equal to 0.

The worksheet A7 of the Excel file, ***A Illustrations.xls,*** *contains 60 months of returns for IBM during the period January 2000 through December 2004. Test the null hypothesis that these mean monthly return equals 0 at the 5% probability level. Also, compute the 95% confidence interval for the mean monthly return for IBM.*

The first test is to compute the mean and standard deviation of the sample of 60 return observations: $\hat{\mu}_X = 0.00408$ and $\hat{\sigma}_X = 0.10378$. Next we compute the standard error of $\hat{\mu}_X$:

$$s_{\hat{\mu}_X} = \hat{\sigma}_X / \sqrt{n} = 0.10378/\sqrt{60} = 0.01340$$

Finally, compute the t-statistic: $t = 0.00408/0.01340 = 0.305$. The OPTVAL library contains a function for computing a t-test of the mean from a pre-specified constant. Its syntax is

OV_STAT_TCNST(*x, cnst, out*)

where x is the vector of sample observations, *cnst* is the prespecified constant, and *out* is an indicator variable instructing the output to be aligned horizontally ("h" or "H") or vertically ("v" or "V"). The output of the function (the t-ratio and the number of degrees of freedom) is written to two adjacent cells, and both must be highlighted when entering the input information. Then press Shift, Ctrl, and Enter simultaneously.

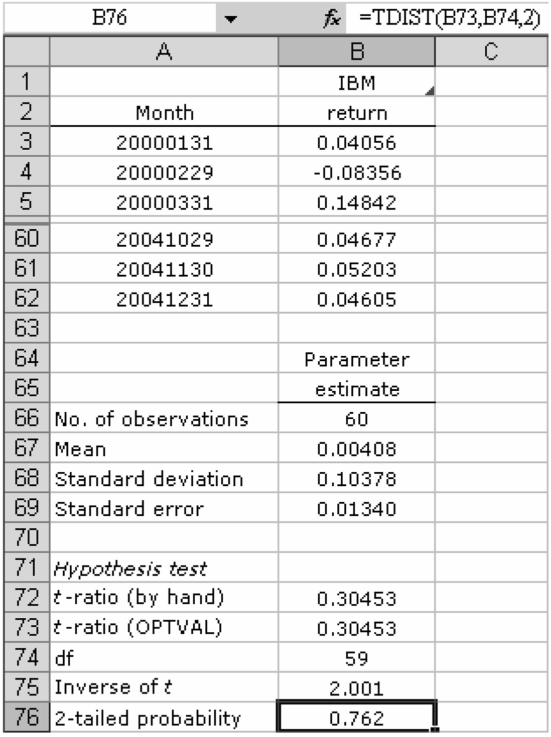
With 59 degrees of freedom and a 5% probability level, the critical t-value is about 2.00. (The critical t-value reported in Table C.3 is 2.000 at 60 degrees of freedom. No value is reported for 59 degrees of freedom). Since the absolute value of 0.305 is less than 2.00, we do not reject the hypothesis that the mean monthly return for IBM is 0. Note that Excel has an add-in function that allows a more accurate value of the critical value. The syntax of the function is

TINV(*probability*, *deg_freedom*)

where *probability* is the desired level of probability in a two-tailed test and *deg_freedom* is the number of degrees of freedom. TINV(0.05, 59) = 2.001, which is very close to our approximate value obtained from Table C.3. Finally, we can use the computed t-ratio directly in the Excel add-in,

TDIST(*x, deg_freedom, tails*)

where x is the t-ratio, *deg_freedom* is the number of degrees of freedom, and *tails* is 1 or 2, depending upon whether you want to perform a one- or two-tailed test. TDIST(0.30453, 59, 2) = 0.762, which means there is a 76.2% probability that the true difference between the population mean and 0 lies outside the range –0.30453 and 0.30453.



B76 ▾ fx =TDIST(B73,B74,2)

	A	B	C
1		IBM	
2	Month	return	
3	20000131	0.04056	
4	20000229	-0.08356	
5	20000331	0.14842	
60	20041029	0.04677	
61	20041130	0.05203	
62	20041231	0.04605	
63			
64		Parameter	
65		estimate	
66	No. of observations	60	
67	Mean	0.00408	
68	Standard deviation	0.10378	
69	Standard error	0.01340	
70			
71	*Hypothesis test*		
72	*t*-ratio (by hand)	0.30453	
73	*t*-ratio (OPTVAL)	0.30453	
74	df	59	
75	Inverse of *t*	2.001	
76	2-tailed probability	0.762	

The 95% confidence interval for the mean of IBM's monthly returns is closely related to the test of the null hypothesis that the mean return equals zero. Substituting the problem parameters into (A.27), we find that

$$\mu_X \leq 0.00408 \pm 2.001(0.01340) = (-0.02273, 0.03089)$$

In other words, based on the 60 months of sample information, we are 95% confident that the "true" mean monthly return of IBM is somewhere between –2.273% and 3.089%—not a high degree of precision indeed. Since 0% is contained within the confidence interval, the null hypothesis that the mean return is 0% cannot be rejected at the 5% level of probability. Similarly, the null hypothesis that the mean monthly return of IBM is 3% cannot be rejected since, it too, falls within the 95% confidence interval.

Test for Equivalence of Means Tests of the equivalence of two means come in two forms. The distinction is driven by the decision about whether it is reasonable to assume the two distributions have the same variance. If two distributions are thought to have the same variance, the appropriate test statistic is

$$t = \frac{\hat{\mu}_X - \hat{\mu}_Y}{\hat{\sigma}_D} \tag{A.32}$$

where

$$\hat{\sigma}_D = \sqrt{\frac{\hat{\sigma}_X^2(n_X - 1) + \hat{\sigma}_Y^2(n_Y - 1)}{n_X + n_Y - 2}\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)} \tag{A.33}$$

We evaluate the significance of this t-value for the Student's distribution with $n_X + n_Y - 2$ degrees of freedom. Note that, if Y is a constant 0, the expression for the standard error becomes

$$\hat{\sigma}_D = \sqrt{\frac{\hat{\sigma}_X^2(n_X - 1)}{n_X - 1}\left(\frac{1}{n_X}\right)} = \hat{\sigma}_X/\sqrt{n_X} = s_{\hat{\mu}_X}$$

which is identical to the standard error in (A.30).

Often there is no reason to believe that the variances of a and b are equal. In this instance, the t-test for the difference in means must be modified. The relevant t-statistic for unequal variance is

$$t = \frac{\hat{\mu}_X - \hat{\mu}_Y}{\sqrt{\hat{\sigma}_X/n_X + \hat{\sigma}_Y/n_Y}} \tag{A.34}$$

where this statistic is *approximately* as Student's t with a number of degrees of freedom equal to

$$df = \frac{(\hat{\sigma}_X^2/n_X + \hat{\sigma}_Y^2/n_Y)^2}{\frac{(\hat{\sigma}_Y^2/n_Y)^2}{n_X - 1} + \frac{(\hat{\sigma}_X^2/n_X)^2}{n_X - 1}} \tag{A.35}$$

Note that expression for determining the number of degrees of freedom (A.35) is, in general, not an integer—there is no reason it has to be.

ILLUSTRATION A.8 Test hypothesis difference in means is 0.

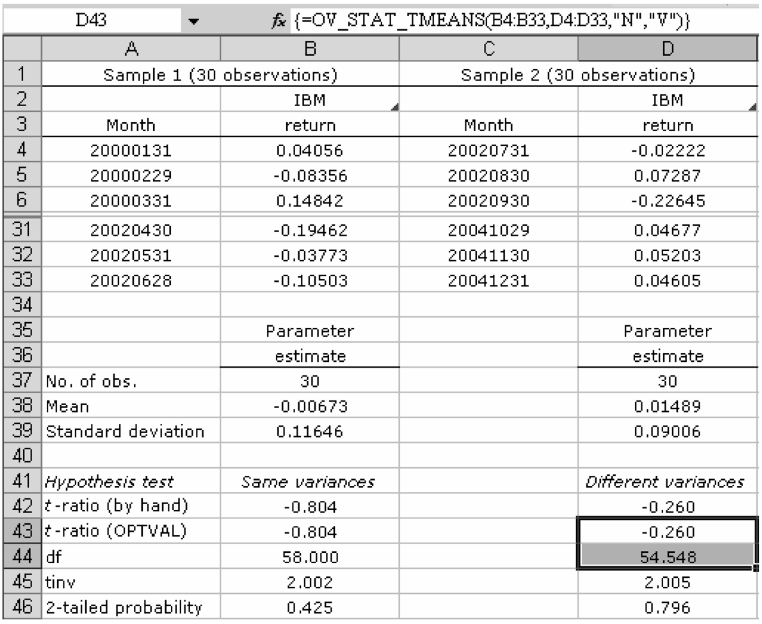
*The worksheet **A8** of the Excel file, **A Illustrations.xls**, contains 60 months of returns for IBM during the period January 2000 through December 2004. Test the null hypothesis that the mean during the first 30 months is no different than the mean return in the second 60 months. First, assume the variances of the two samples are equal, and then assume the variances are different.*

After computing the mean and variance of each sample, we can perform the computations by hand using equations (A.32) through (A.35). But both computations can also be performed using the OPTVAL function

OV_STAT_TMEANS(*x*, *y*, *ind*, *out*)

where x and y are the vectors of sample observations for the two samples, *ind* is an indicator variable instructing the function to assume equal variances ("y" or "Y") or unequal variances ("n" or "N"), and *out* is an indicator variable instructing the function to return the output horizontally ("h" or "H") or vertically ("v" or "V"). Again, the output is the t-ratio and the number of degrees of freedom and so two adjacent cells must be highlighted when the function is called. The results are shown below.

The results indicate that there is little reason to believe that (1) the mean return for IBM is different in the two sample periods; and (2) different variances have an important effect on the testing procedure. Under the assumption that the variances are the same across samples, the t-ratio for testing the null hypothesis that the means are the same is –0.804. Since the critical value of the t-distribution corresponding to a two-tailed test and 58 degrees of freedom is $t_{0.05/2,58} = 2.002$. Since the absolute value of the t-ratio is less than 2.002, we cannot reject the hypothesis that the means are the same. Alternatively, since the p-value, 0.425, is greater than the demanded level of significance, 0.05, the null cannot be rejected.



D43 fx {=OV_STAT_TMEANS(B4:B33,D4:D33,"N","V")}

	A	B	C	D
1	Sample 1 (30 observations)		Sample 2 (30 observations)	
2		IBM		IBM
3	Month	return	Month	return
4	20000131	0.04056	20020731	-0.02222
5	20000229	-0.08356	20020830	0.07287
6	20000331	0.14842	20020930	-0.22645
31	20020430	-0.19462	20041029	0.04677
32	20020531	-0.03773	20041130	0.05203
33	20020628	-0.10503	20041231	0.04605
34				
35		Parameter		Parameter
36		estimate		estimate
37	No. of obs.	30		30
38	Mean	-0.00673		0.01489
39	Standard deviation	0.11646		0.09006
40				
41	*Hypothesis test*	*Same variances*		*Different variances*
42	*t*-ratio (by hand)	-0.804		-0.260
43	*t*-ratio (OPTVAL)	-0.804		-0.260
44	df	58.000		54.548
45	tinv	2.002		2.005
46	2-tailed probability	0.425		0.796

Test for Equivalence of Means in a Paired Sample Paired comparisons in finance-related problems are not infrequent. Suppose, for example, two stocks have done particularly well during a specified period of time, but that, during the same period, the stock market did particularly well. Is the performance of the two stocks different in a meaningful way?

To answer this question, we can, again, rely on a t-test. The t-ratio is

$$t = \frac{\hat{\mu}_X - \hat{\mu}_Y}{\hat{\sigma}_D} \tag{A.36}$$

and is evaluated with degrees of freedom. The definition of the denominator is

$$\hat{\sigma}_D = \left(\frac{\hat{\sigma}_X^2 + \hat{\sigma}_Y^2 - 2\hat{\sigma}_{XY}}{n-1} \right)^{0.5} \tag{A.37}$$

that is, the standard error of the difference in returns of *X* and *Y*. A little reflection will tell you why this is appropriate. Since both *a* and *b* may co-vary with some factor, we need to abstract from that factor. Thus, we reduce the variance in the numerator of (A.37) by the amount of the covariation in determining whether the difference is indeed significant.

ILLUSTRATION A.9 Test hypothesis difference between means in paired sample.

*The worksheet **A9** of the Excel file, **A Illustrations.xls,** contains 60 months of returns for IBM and GM during the sample period January 2000 through December 2004. Test the null hypothesis that the mean of IBM's returns is different from the mean of GM's returns.*

Summary statistics for the return series are shown in the table below. Since individual stock returns tend to covary with the market, they tend to covary with each other. To check if this is the case, we can compute the correlation between the return series. The estimated correlation coefficient is 0.294, which indicates that, when testing for a difference between the mean returns of the two stocks, it is appropriate to use a test statistic that accounts for the contemporaneous relation between the series.

With the information provided in the summary table, we can compute the *t*-ratio using (A.36) and (A.37). The *t*-ratio is 0.300. Using a two-tailed test with 59 degrees of freedom and $\alpha = 0.05$, we cannot reject the hypothesis that the mean returns of IBM and GM are the same. The OPTVAL library contains a function for computing the *t*-ratio directly without us having to perform the intermediate computations. Its syntax is

OV_STAT_TPMEANS(*x*, *y*, *out*)

where *x* and *y* are the vectors containing the pairs of observations vectors, and *out* is an indicator variable instructing the function to return the output horizontally ("h" or "H") or vertically ("v" or "V") . The output of the function is the *t*-ratio and the number of degrees of freedom.

	Parameter Estimates	
No. of obs.	60	60
Mean	0.00408	–0.00078
Standard deviation	0.10378	0.10721
Variance	0.01077	0.01149
Covariance	0.00321	

Correlation Matrix	**IBM**	**GM**
IBM	1	
GM	0.294	1

Hypothesis Test	
t-ratio	0.300
df	59
tinv	2.001
2-tailed probability	0.765

Type I and Type II Errors and the Power of a Test

With the rules for conducting hypothesis tests and building confidence intervals in hand, we are in a position to discuss two more subtle statistical issues. The first is Type I and Type II errors. Recall that, in the illustrations of this appendix, we preset the desired level of significance of the test *before* the test was performed. The choice of the level of significance α (usually 5% or 1%) represents the probability of rejecting the null hypothesis when the null hypothesis is, in fact, true. This type of mistake is called *Type I error. Type II error*, on the other hand, refers to the probability that the null hypothesis is not refuted when it should be.

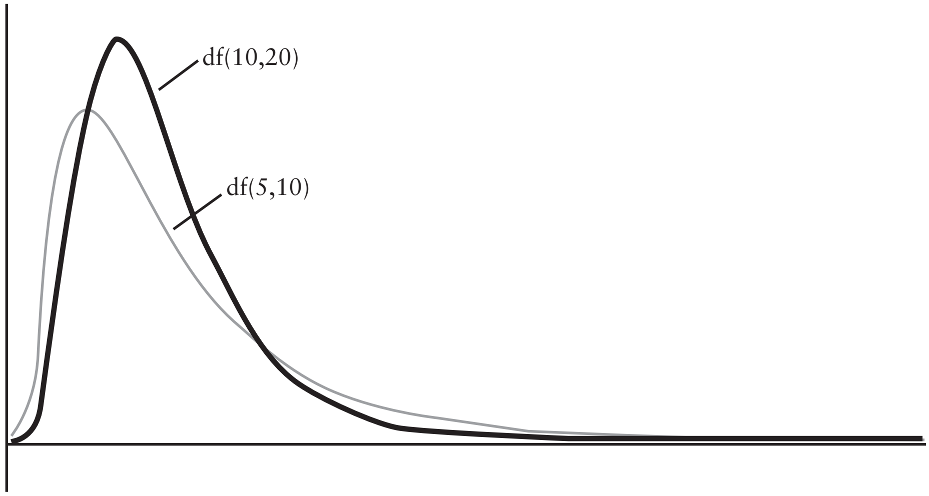
To more clearly distinguish between the two types of errors, consider changing the level of significance in a test from 5% to 1%. Obviously, the probability of incorrectly rejecting the null hypothesis (Type I error) falls from 5% to 1%. At the same time, the probability of a Type II error increases. The lower the value of α, the wider the range of outcomes within the confidence interval, and the greater our inability to distinguish between values contained within the interval. If the true population parameter is 3 and the confidence interval is (–5,+5), a significance test will not reject the null hypothesis that the parameter is 0, even though we know that it is not. Thus, in selecting the level of significance, we face a trade off. As we lower the probability of Type I error, we increase the probability of Type II error. The choice between the two types of errors depends on the particular problem. In finance applications, we usually choose a low level of significance and, hence, a low probability of Type I error.

Closely related to Type I and Type II errors is the concept of the power of a test. Suppose that we fail to reject the hypothesis that the population parameter is 0. Consider the possible reasons for this "failure." One obvious reason is that the null hypothesis is true. Another possibility is that the null hypothesis is false, but the particular data set used for the test happens to be consistent with the null. The statistical concept that helps us evaluate the importance of the second explanation is the *power of a test. Power* is the probability of rejecting the null hypothesis when it is in fact false and is, therefore, equal to one minus the probability of a Type II error (i.e., one minus the probability that one will accept the null hypothesis as true when it is in fact false). Note that power depends not only on the size of the effect that has been measured, but also on the number of observations in the sample. Holding other factors constant, the larger the effect and the larger the sample size, the more powerful the test. When a statistical analysis with relatively low power fails to show a significant p-value, we should not be hasty in concluding that there is no effect. We must allow for the fact that the study may be inconclusive because the data set is not rich enough sufficient to allow us to distinguish between the null and alternative hypotheses.

F Distribution

Formally defined, $(X/n_1)/(Y/n_2)$ is distributed according to an *F distribution* with n_1 and n_2 degrees of freedom if X and Y are independent and distributed as chi square with n_1 and n_2 degrees of freedom, respectively. The F-distribution is skewed to the right, as shown in Figure A.8. The exact shape will depend on the numbers of degrees of freedom in the numerator and the denominator. The fig-

FIGURE A.8 *F*-distribution with (5,10) and (10,20) degrees of freedom.



ure displays an *F*-distribution with 5 and 10 degrees of freedom and another with 10 and 20 degrees of freedom. The latter distribution is less skewed.

Test for Equivalence of Variances The *F*-distribution is commonly used in tests of the equality of two variances. The *F*-statistic is always tabulated with the larger estimate of variance in the numerator and the smaller estimate in the denominator. Thus assuming

$$\hat{\sigma}_X^2 > \hat{\sigma}_Y^2$$

the *F*-statistic is

$$F = \frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2} \tag{A.38}$$

with $n_X - 1$ and $n_Y - 1$ degrees of freedom. The resulting ratio is always greater than 1, and provides information about the upper tail of the *F*-distribution. The greater the difference between the two variances, the greater the *F*-statistic. Thus, a large value of *F* implies that it is unlikely that the two error variances are equal. Tables C.4A and C.4B summarize critical *F*-values under 5% and 1% probability levels. Note that the tables are arranged with the columns representing different numbers of degrees of freedom in the numerator and the rows representing different numbers of degrees of freedom in the denominator. To illustrate applying the tables, assume the number of degrees of freedom in the numerator and the denominator is 10 and that we preset the level of significance to $\alpha = 0.05$. The critical F-value in Table C.4A is 2.98, which means that if the F-statistic from the test exceeds 2.98, we reject the null hypothesis that the two variances are equal at the 0.05 probability level. If the F-statistic exceeds 4.85, we also reject the null hypothesis that the two variances are equal at the 0.01 probability level (see Table C.4B).

ILLUSTRATION A.10 Test for difference in variances of stock return series.

The worksheet ***A10*** *of the Excel file,* ***A Illustrations.xls,*** *contains 60 months of returns for IBM and the market portfolio during the sample period January 2000 through December 2004. Test the null hypothesis that the variance of IBM's returns is different from the variance of the returns of the market at the .05 probability level.*

Summary statistics for the return series are shown in the table below. The variance of IBM returns is considerably larger than the variance of the market returns, so we place IBM in the numerator. The F-statistic is

$$F = \frac{0.01077/(60-1)}{0.00242/(60-1)} = 4.4432$$

With $\alpha = 0.05$ and 59 degrees of freedom in both the numerator and the denominator, the critical value $F_{0.05,59,59}$ is 1.5400. The closest value in Table C.4A is $F_{0.05,60,60} = 1.53$. The exact value was obtained using the Excel function

FINV(*probability,deg_freedom1,deg_freedom2*)

where *probability* is the preset significance level, and *deg_freedom1* and *deg_freedom2* are the number of degrees of freedom in the numerator and denominator, respectively. FINV(0.05,59,59) = 1.5400. Finally, Excel also has a function for computing the p-value of an F-statistic directly. Its syntax is

FDIST (*x,deg_freedom1,deg_freedom2*)

where x is the sample F-statistic. As it turns out, FDIST(4.4432,59,59) = 0.0000. The null hypothesis that the variances of the two series are equal is soundly rejected.

Month	IBM Return	Market Return
20000131	0.04056	–0.03977
20000229	–0.08356	0.03178
20000331	0.14842	0.05353
20041029	0.04677	0.01780
20041130	0.05203	0.04826
20041231	0.04605	0.03518

	Parameter Estimates	
No. of obs.	60	60
Mean	0.00408	0.00018
Standard deviation	0.10378	0.04924
Variance	0.01077	0.00242

Hypothesis Test

F-statistic (by hand)	4.4432	
df(num,den)	59	59
finv	1.5400	
Probability	0.0000	

Test for Autocorrelation

In the study of finance, another key property of the returns (besides normality) is *independence*—past returns carry no information regarding current and future returns. The usual way of testing for whether returns are independently distributed is by calculating the *sample autocorrelation function*

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (X_t - \overline{X})(X_{t+k} - \overline{X})}{\sum_{t=1}^{T-k} (X_t - \overline{X})^2} \tag{A.39}$$

where T is the number of observations in the time series. If the returns are independent, the lag k autocorrelation should be zero. To test whether a particular value of the autocorrelation function ρ_k is equal to zero, we use a Bartlett test. Under the null hypothesis that the time series is *white noise,* the sample autocorrelation coefficients are approximately normally distributed with mean zero and standard deviation $1/\sqrt{T}$. For the S&P 500 monthly returns in our sample, the autocorrelation function is:

Lag	1	2	3	4	5
Autocorrelation	–0.0785	–0.0409	0.0314	–0.0570	–0.0180
Standard deviation	0.0783	0.0783	0.0783	0.0783	0.0783

It is computed using the OPTVAL function

OV_STAT_AUTOCORREL(*k*, *x*, *out*)

where *k* is the maximum number of lags, *x* is the time series, and *out* is an indicator variable set equal to 0 if the output array is to be returned horizontally and 1 if the array is to be returned vertically. With 163 monthly returns in the time series, the standard error is $1/\sqrt{163} = 0.07833$. In other words, the absolute magnitude of an autocorrelation coefficient would have to be greater than $0.07833 \times 2 = 0.15665$ in order to sure that the autocorrelation coefficient is not zero with 95% confidence. The sample autocorrelation function indicates that none of the true coefficients are different from zero. Box and Pierce developed a Q statistic for testing the *joint* hypothesis that *all* the autocorrelation coefficients are zero, that is,

$$Q = T \sum_{k=1}^{K} \hat{\rho}_k^2 \tag{A.40}$$

where Q is (approximately) distributed as chi-square with K degrees of freedom. The OPTVAL function

OV_STAT_BOX_PIERCE (*k, x, out*)

computes the chi-squared statistics for different values of *k*. The results for the S&P 500 monthly returns are

Lag	1	2	3	4	5
Autocorrelation	–0.0785	–0.0409	0.0314	–0.0570	–0.0180
Standard deviation	0.0783	0.0783	0.0783	0.0783	0.0783
Box-Pierce Q statistic	1.0051	1.2778	1.4386	1.9675	2.0205
Critical chi-square level	2.7055	4.6052	6.2514	7.7794	9.2364

The levels of the Q-statistic are well below their critical levels at the 90% confidence level, so we cannot reject the hypothesis that all the *true* autocorrelation coefficients are equal to zero.

Central Limit Theorem

Earlier we stated that the parameters of the distribution are certain if the entire population is known (i.e., $n = N$). Intuitively, therefore, this must mean that, as the sample size grows large, the estimate of the mean should converge on the population mean. This intuition, which holds for probability distributions with finite means, can be summed up formally as:

> **The central limit theorem.** If the random variable X has mean μ_X and variance σ_X^2, then the sampling distribution of $\hat{\mu}_X$ becomes approximately normal with mean μ_X and variance σ_X^2/n as n increases.

In other words, for sufficiently large sample sizes, we can rely on the normality assumption, which greatly simplifies statistical tests. The central limit theorem will prove useful in assessing the performance of option trading strategies in Chapter 10.

REFERENCES AND SUGGESTED READINGS

Abramowitz, Milton, and Irene A. Stegum. 1972. *Handbook of Mathematical Functions*, 10th ed. Washington, DC: National Bureau of Standards.

Bartlett, M. S. 1946. On the theoretical specification of sampling properties of autocorrelated time series. *Journal of the Royal Statistical Society, Series B8* 27.

Intriligator, Michael D. 1978. *Econometric Techniques, and Applications*. Englewood Cliffs, NJ: Prenctice-Hall.

Jarque, C. M., and A. K. Bera. 1987. A test of normality of observations and regression residuals, *International Statistical Review* 11: 351–360.

Jarque, C. M., and A. K. Bera. 1980. Efficient tests of normality, homoscedasticity and serial dependence of regression residuals. *Economic Letters* 6, 255–259.

Joanes, D. N., and C. A. Gill. 1998. Comparing measures of sample skewness and kurtosis. *Journal of the Royal Statistical Society (Series D): The Statistician* 47: 183–189 .

Jorion, Philippe. 1997. *Value at Risk*. Homewood, IL: Irwin.

Kennedy, Peter. 1992. *A Guide to Econometrics*, 3rd ed. Cambridge, MA: MIT Press.

Kmenta, Jan. 1971. *Elements of Econometrics*. New York: Macmillan.

Pindyck, Robert S., and Daniel L. Rubinfeld. 1998. *Econometric Models and Economic Forecasts*, 4th ed. Boston: Irwin/McGraw-Hill.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge: Cambridge University Press.

Student (W. S. Gosset). 1908. The probable error of a mean. *Biometrika* 6 (1): 1–25.