EAY and r_{MM} are two ways to annualize an HPY. Different instruments have different conventions for quoting yields. In order to compare the yields on instruments with different yield conventions, you must be able to convert the yields to a common measure. For instance, to compare a T-bill yield and a LIBOR yield, you can convert the T-bill yield from a bank discount yield to a money market yield and compare it to the LIBOR yield (which is already a money market yield). In order to compare yields on other instruments to the yield (to maturity) of a semi-annual pay bond, we simply calculate the effective semiannual yield and double it. A yield calculated in this manner is referred to as a *bond equivalent yield* (BEY).

Statistical Concepts and Market Returns

Cross-Reference to CFA Institute Assigned Reading #8

The two key areas you should concentrate on in this reading are measures of central tendency and measures of dispersion. Measures of central tendency include the arithmetic mean, geometric mean, weighted mean, median, and mode. Measures of dispersion include the range, mean absolute deviation, variance, and standard deviation. When describing investments, measures of central tendency provide an indication of an investment's expected value or return. Measures of dispersion indicate the riskiness of an investment (the uncertainty about its future returns or cash flows).

Measures of Central Tendency

Arithmetic mean. A population average is called the population mean (denoted μ). The average of a sample (subset of a population) is called the sample mean (denoted $\bar{x}$). Both the population and sample means are calculated as arithmetic means (simple average). We use the sample mean as a "best guess" approximation of the population mean.

Median. Middle value of a data set, half above and half below. With an even number of observations, median is the average of the two middle observations.

Mode. Value occurring most frequently in a data set. Data set can have more than one mode (bimodal, trimodal, etc.) but only one mean and one median.

Geometric mean:

- Used to calculate compound growth rates.
- If returns are constant over time, geometric mean equals arithmetic mean.
- The greater the variability of returns over time, the greater the difference between arithmetic and geometric mean (arithmetic will always be higher).

©2017 Kaplan, Inc.

- When calculating the geometric mean for a returns series, it is necessary to add one to each value under the radical, and then subtract one from the result.
- The geometric mean is used to calculate the time-weighted return, a performance measure.

$$\text{geometric mean return} = R_G = \sqrt[n]{(1+R_1)\times(1+R_2)\times\ldots\times(1+R_n)}-1$$

Example:

A mutual fund had the following returns for the past three years: 15%, –9%, and 13%. What is the arithmetic mean return, the 3-year holding period return, and the average annual compound (geometric mean) return?

Answer:

$$\text{arithmetic mean: } \frac{15\% - 9\% + 13\%}{3} = 6.333\%$$

$$\text{holding period return: } 1.15 \times 0.91 \times 1.13 - 1 = 0.183 = 18.3\%$$

$$\begin{aligned}\text{geometric mean: } R_G &= \sqrt[3]{(1+0.15)\times(1-0.09)\times(1+0.13)}-1\\ &= \sqrt[3]{1.183}-1 = 1.0575-1 = 0.0575 = 5.75\%\end{aligned}$$

Geometric mean return is useful for finding the yield on a zero-coupon bond with a maturity of several years or for finding the average annual growth rate of a company's dividend or earnings across several years. Geometric mean returns are a compound return measure.

Weighted mean. Mean in which different observations are given different proportional influence on the mean:

$$\text{weighted mean} = \overline{X}_w = \sum_{i=1}^{n} w_i X_i = (w_1 X_1 + w_2 X_2 + \ldots + w_n X_n)$$

where:
$X_1, X_2, \ldots, X_n$ = observed values
$w_1, w_2, \ldots, w_n$ = corresponding weights for each observation, $\sum w_i = 1$

©2017 Kaplan, Inc.

Weighted means are used to calculate the actual or expected return on a portfolio, given the actual or expected returns for each portfolio asset (or asset class). For portfolio returns, the weights in the formula are the percentages of the total portfolio value invested in each asset (or asset class).

Example: Portfolio return

A portfolio is 20% invested in Stock A, 30% invested in Stock B, and 50% invested in Stock C. Stocks A, B, and C experienced returns of 10%, 15%, and 3%, respectively. Calculate the portfolio return.

Answer:

$$R_p = 0.2(10\%) + 0.3(15\%) + 0.5(3\%) = 8.0\%$$

A weighted mean is also used to calculate the expected return given a probability model. In that case, the weights are simply the probabilities of each outcome.

Example: Expected portfolio return

A portfolio of stocks has a 15% probability of achieving a 35% return, a 25% chance of achieving a 15% return, and a 60% chance of achieving a 10% return. Calculate the expected portfolio return.

Answer:

$$E(R_p) = 0.15(35) + 0.25(15) + 0.60(10) = 5.25 + 3.75 + 6 = 15\%$$

Note that an arithmetic mean is a weighted mean in which all of the weights are equal to 1/n (where *n* is the number of observations).

Measures of Dispersion

Range is the difference between the largest and smallest value in a data set and is the simplest measure of dispersion. You can think of the dispersion as measuring the width of the distribution. The narrower the range, the less dispersion.

For a population, *variance* is defined as the average of the squared deviations from the mean.

 ©2017 Kaplan, Inc.

Example:

Stocks A, B, and C had returns of 10%, 30%, and 20%, respectively. Calculate the population variance (denoted σ^2) and sample variance (denoted s^2).

Answer:

The process begins the same for population and sample variance.

Step 1: Calculate the mean expected return: $\frac{(10+30+20)}{3} = 20$

Step 2: Calculate the squared deviations from the mean and add them together: $(10-20)^2 + (30-20)^2 + (20-20)^2 = 100 + 100 + 0 = 200$

Step 3: Divide by number of observations (n = 3) for the population variance and by the number of observations minus one for the sample variance:

$$\text{population variance} = \sigma^2 = \frac{200}{3} = 66.67$$

$$\text{sample variance} = s^2 = \frac{200}{3-1} = \frac{200}{2} = 100$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N} \qquad s^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}$$

Standard deviation is the square root of variance. On the exam, if the question is asking for the standard deviation, do not forget to take the square root!

Coefficient of variation expresses how much dispersion exists relative to the mean of a distribution and allows for direct comparison of the degree of dispersion across different data sets. It measures risk per unit of expected return.

$$CV = \frac{\text{standard deviation of returns}}{\text{mean return}}$$

When comparing two investments using the CV criterion, the one with the lower CV is the better choice.

The *Sharpe ratio* is widely used to evaluate investment performance and measures excess return per unit of risk. Portfolios with large Sharpe ratios are preferred to portfolios with smaller ratios because it is assumed that rational investors prefer higher excess returns (returns in excess of the risk-free rate) and dislike risk.

$$\text{Sharpe ratio} = \frac{\text{excess return}}{\text{risk}} = \frac{R_{\text{portfolio}} - R_{\text{risk-free}}}{\sigma_p}$$

If you are given the inputs for the Sharpe ratio for two portfolios and asked to select the best portfolio, calculate the Sharpe ratio, and choose the portfolio with the higher ratio.

Skewness and Kurtosis

Skewness represents the extent to which a distribution is not symmetrical.

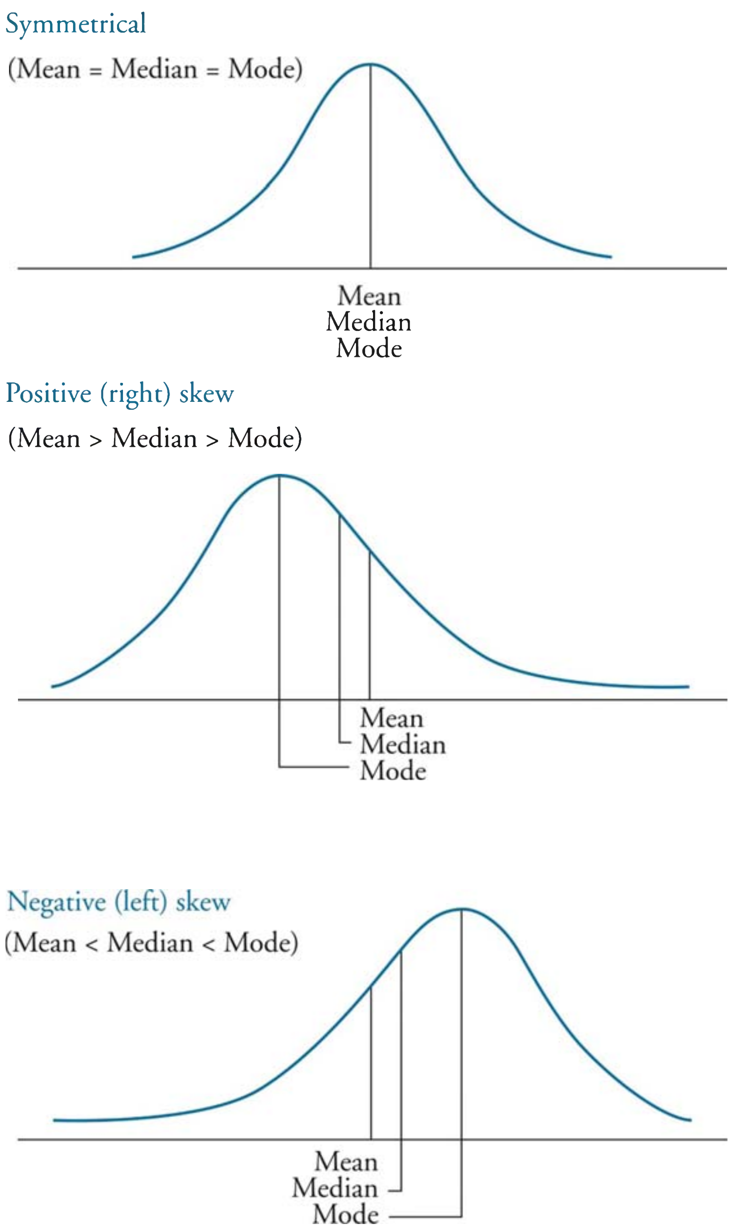
A *right-skewed* distribution has positive skew (or skewness) and a mean that is greater than the median, which is greater than the mode.

A *left-skewed* distribution has negative skewness and a mean that is less than the median, which is less than the mode.

The attributes of normal and skewed distributions are summarized in the following illustration.

©2017 Kaplan, Inc.

Figure 1: Skewed Distributions



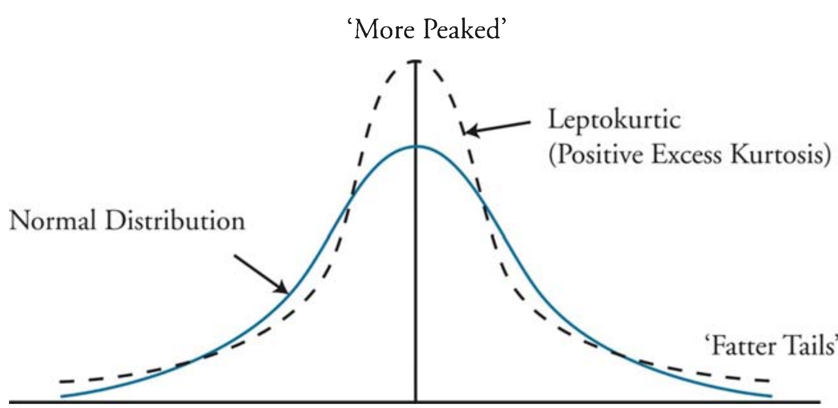
To remember the relations, think of "pulling on the end" of a normal distribution, which is symmetrical with the mean, median, and mode equal. If you pull on the right or positive end, you get a right-skewed (positively skewed) distribution. If you can remember that adding extreme values at one end of the distribution has the greatest effect on the mean, and doesn't affect the mode or high point of the distribution, you can remember the relations illustrated in the preceding graph.

Kurtosis is a measure of the degree to which a distribution is more or less peaked than a normal distribution, which has kurtosis of 3.

Excess kurtosis is kurtosis relative to that of a normal distribution. A distribution with kurtosis of 4 has excess kurtosis of 1. It is said to have positive excess kurtosis. A distribution with positive excess kurtosis (a leptokurtic distribution) will have more returns clustered around the mean and more returns with large deviations from the mean (fatter tails). In finance, positive excess kurtosis is a significant issue in risk assessment and management, because fatter tails means an increased probability of extreme outcomes, which translates into greater risk.

An illustration of the shapes of normal and leptokurtic distribution is given in the following graph.

Figure 2: Kurtosis



Probability Concepts

Cross-Reference to CFA Institute Assigned Reading #9

The ability to apply probability rules is important for the exam. Be able to calculate and interpret widely used measures such as expected value, standard deviation, covariance, and correlation.

Important Terms

- *Random variable*. Uncertain quantity/number.
- *Outcome*. Realization of a random variable.
- *Event*. Single outcome or a set of outcomes.
- *Mutually exclusive events*. Cannot both happen at same time.
- *Exhaustive set of events*. Set that includes all possible outcomes.

The probability of any single outcome or event must not be less than zero (will not occur) and must not be greater than one (will occur with certainty). A *probability function* (for a discrete probability distribution) defines the probabilities that each outcome will occur. To have a valid probability function, it must be the case that

©2017 Kaplan, Inc.

the sum of the probabilities of any set of outcomes or events that is both mutually exclusive and exhaustive is 1 (it is certain that a random variable will take on one of its possible values). An example of a valid probability function is:

$$\text{Prob}(x) = x/15 \text{ for possible outcomes, } x = 1, 2, 3, 4, 5$$

Odds For and Against

If the probability of an event is 20%, it will occur, on average, one out of five times. The "odds for" are 1-to-4 and the "odds against" are 4-to-1.

Multiplication Rule for Joint Probability

$$P(AB) = P(A \mid B) \times P(B) = P(B \mid A) \times P(A)$$

The probability that A and B will both (jointly) occur is the probability of A given that B occurs, multiplied by the (unconditional) probability that B will occur.

Addition Rule

$$P(A \text{ or } B) = P(A) + P(B) - P(AB)$$

If A and B are mutually exclusive, $P(AB)$ is zero and $P(A \text{ or } B) = P(A) + P(B)$

Used to calculate the probability that at least one (one or both) of two events will occur.

Total Probability Rule

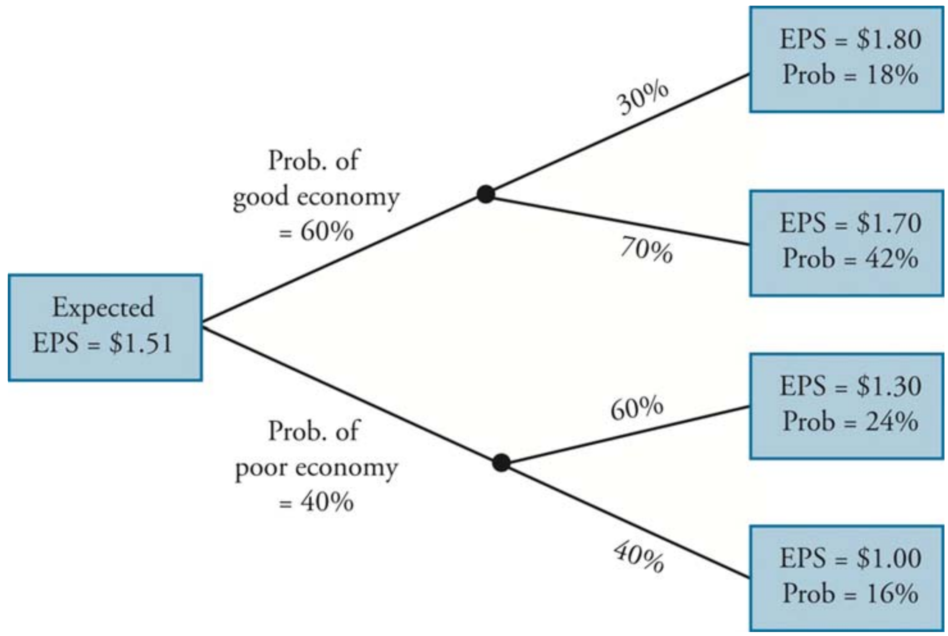
$$P(R) = P(R \mid I) \times P(I) + P(R \mid I^C) \times P(I^C)$$

where: I and I^C are *mutually exclusive and an exhaustive set of events* (i.e., if I occurs, then I^C cannot occur and one of the two must occur).

A tree diagram shows a variety of possible outcomes for a random variable, such as an asset price or earnings per share.

©2017 Kaplan, Inc.

Figure 3: A Tree Diagram for an Investment Problem



We can illustrate several probability concepts with a tree diagram. The (unconditional) expected EPS is the sum of the possible outcomes, weighted by their probabilities.

$$0.18 \times 1.80 + 0.42 \times 1.70 + 0.24 \times 1.30 + 0.16 \times 1.00 = \$1.51$$

The (conditional) expectation of EPS, given that the economy is good, is $1.73 = 0.3(1.80) + 0.7(1.70). Expected EPS, given that the economy is poor, is 0.6(1.30) + 0.4(1.00) = $1.18.

The probabilities of each of the EPS outcomes are simply the product of the two probabilities along the (branches) of the tree [e.g., P(EPS = $1.80) = 0.6 × 0.3 = 18%].

Covariance

The *covariance* between two variables is a measure of the degree to which the two variables tend to move together. It captures the linear relationship between one random variable and another.

A *positive covariance* indicates that the variables tend to move together; a *negative covariance* indicates that the variables tend to move in opposite directions relative

©2017 Kaplan, Inc.

to their means. Covariance indicates the direction of the relationship and does not directly indicate the strength of the relationship. Therefore, if you compare the covariance measures for two sets of (paired) random variables and the second is twice the value of the first, the relationship of the second set isn't necessarily twice as strong as the first because the variance of the variables may be quite different as well.

Example:

Covariance can be calculated using a joint probability table as follows:

	$R_X = 15\%$	$R_X = 10\%$
$R_Y = 20\%$	0.30	0
$R_Y = 5\%$	0	0.70

First, find the expected returns on X and Y:

$$E(R_X) = 0.30(15) + 0.70(10) = 11.5\%$$
$$E(R_Y) = 0.30(20) + 0.70(5) = 9.5\%$$

Next calculate the covariance:

$$\begin{aligned} Cov(R_X, R_Y) &= [0.3(15.0 - 11.5)(20.0 - 9.5)] + [0.7(10.0 - 11.5)(5.0 - 9.5)] \\ &= 11.025 + 4.725 = 15.75 \end{aligned}$$

Correlation

The *correlation coefficient, r,* is a standardized measure (unlike covariances) of the strength of the linear relationship between two variables. The correlation coefficient can range from –1 to +1.

$$r = corr(R_i, R_j) = \frac{Cov(R_i, R_j)}{\sigma(R_i)\sigma(R_j)}$$

A correlation of +1 indicates a perfect positive correlation. In that case, knowing the outcome of one random variable would allow you to predict the outcome of the other with certainty.

Expected Return and Variance of a Portfolio of Two Stocks

Know how to compute the *expected return and variance for a portfolio of two assets* using the following formulas:

$$E(R_p) = w_A R_A + w_B R_B$$

$$Var_p = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2 w_A w_B \sigma_A \sigma_B \rho_{A,B}$$

$$Var_p = w_A^2 \sigma_A^2 + w_B^2 \sigma_B^2 + 2 w_A w_B Cov_{A,B}$$

Note that $\sigma_A \sigma_B \rho_{A,B} = Cov_{A,B}$ so the formula for variance can be written either way.

Study Session 3: Quantitative Methods—Applications

Common Probability Distributions

Cross-Reference to CFA Institute Assigned Reading #10

Critical topics to understand include the normal distribution and areas under the normal curve, the *t*-distribution, skewness, kurtosis, and the binomial distribution. Be able to calculate confidence intervals for population means based on the normal distribution.

Discrete random variable: A limited (finite) number of possible outcomes and each has a positive probability. They can be counted (e.g., number of days without rain during a month).

Continuous random variable: An infinite number of possible outcomes. The number of inches of rain over a month can take on an infinite number of values, assuming we can measure it with infinite precision. For a continuous random variable, the probability that the random variable will take on any single one (of the infinite number) of the possible values is zero.

Probability function, p(x), specifies the probability that a random variable equals a particular value, *x*.

A *cumulative density function* (CDF), for either a discrete or continuous distribution, gives the probability that a random variable will take on a value *less than or equal to* a specific value, that is, the probability that the value will be between minus infinity and the specified value.

©2017 Kaplan, Inc.

For the function, Prob(x) = x/15 for x = 1, 2, 3, 4, 5, the CDF is:

$$\sum_{x=1}^{X} \frac{x}{15}, \text{ so that F(3) or Prob } (x \leq 3) \text{ is } 1/15 + 2/15 + 3/15 = 6/15 \text{ or } 40\%$$

This is simply the sum of the probabilities of 1, 2, and 3. Note that
Prob (x = 3, 4) can be calculated as F(4) – F(2) = $\frac{10}{15} - \frac{3}{15} = \frac{7}{15}$.

Uniform Distributions

With a uniform distribution, the probabilities of the outcomes can be thought of as equal. They are equal for all possible outcomes with a discrete uniform distribution, and equal for equal-sized ranges of a uniform continuous distribution.

For example, consider the *discrete uniform probability distribution* defined as X = {1, 2, 3, 4, 5}, p(x) = 0.2. Here, the probability for each outcome is equal to 0.2 [i.e., p(1) = p(2) = p(3) = p(4) = p(5) = 0.2]. Also, the cumulative distribution function for the *n*th outcome, $F(x_n) = np(x)$, and the probability for a range of outcomes is p(x)k, where *k* is the number of possible outcomes in the range.

A *continuous uniform distribution* over the range of 1 to 5 results in a 25% probability [1 / (5 – 1)] that the random variable will take on a value between 1 and 2, 2 and 3, 3 and 4, or 4 and 5, since 1 is one-quarter of the total range of the random variable.

The Binomial Distribution

A **binomial random variable** may be defined as the number of "successes" in a given number of trials where the outcome can be either "success" or "failure." You can recognize problems based on a binomial distribution from the fact that there are only two possible outcomes (e.g., the probability that a stock index will rise over a day's trading). The probability of success, *p*, is constant for each trial, the trials are independent, and the probability of failure (no success) is simply 1 – p. A binomial distribution is used to calculate the number of successes in *n* trials. The probability of *x* successes in *n* trials is:

$$p(x) = P(X = x) = ({}_nC_r)p^x(1 - p)^{n-x}$$

and the expected number of successes is *np*.

If the probability of a stock index increasing each day (*p*) is 60%, the probability (assuming independence) that the index will increase on exactly three of the next five days (and not increase on two days) is $({}_5C_3)0.6^3(1 - 0.6)^2 = 0.3456$.

©2017 Kaplan, Inc.

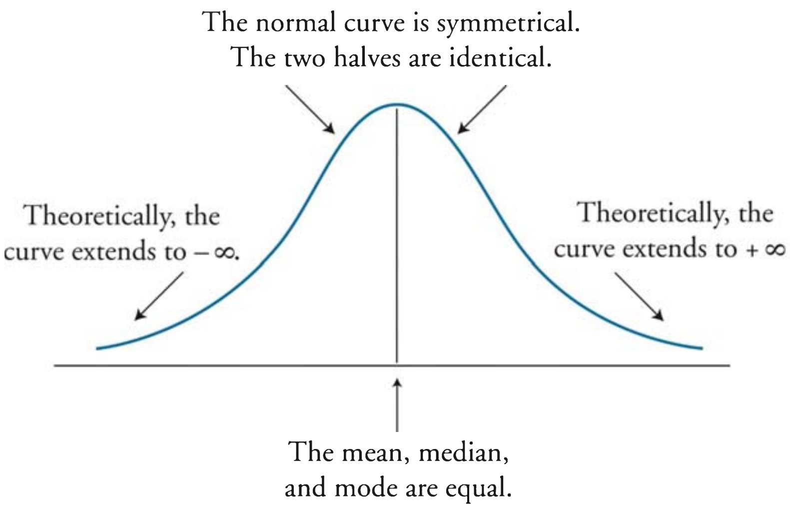
A binomial tree to describe possible stock price movement for *n* periods shows the probabilities for each possible number of successes over *n* periods. Additionally, assuming that the stock price over any single period will either increase by a factor *U* or decrease by a factor 1/U, a binomial tree shows the possible *n*-period outcomes for the stock price and the probabilities that each will occur.

Normal Distribution: Properties

- Completely described by mean and variance.
- Symmetric about the mean (skewness = 0).
- Kurtosis (a measure of peakedness) = 3.
- Linear combination of jointly, normally distributed random variables is also normally distributed.

Many properties of the normal distribution are evident from examining the graph of a normal distribution's probability density function:

Figure 4: Normal Distribution Probability Density Function



Calculating Probabilities Using the Standard Normal Distribution

The ***z-value*** "standardizes" an observation from a normal distribution and represents the number of standard deviations a given observation is from the population mean.

$$z = \frac{\text{observation} - \text{population mean}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

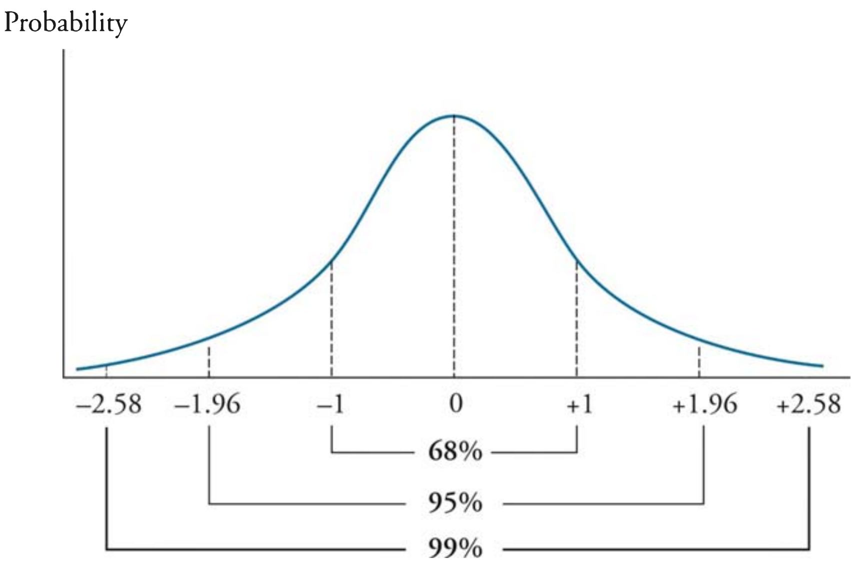
 ©2017 Kaplan, Inc.

Confidence Intervals: Normal Distribution

A *confidence interval* is a range of values around an expected outcome within which we expect the actual outcome to occur some specified percentage of the time.

The following graph illustrates confidence intervals for a standard normal distribution, which has a mean of 0 and a standard deviation of 1. We can interpret the values on the x-axis as the number of standard deviations from the mean. Thus, for any normal distribution we can say, for example, that 68% of the outcomes will be within one standard deviation of the mean. This would be referred to as a 68% confidence interval.

Figure 5: The Standard Normal Distribution and Confidence Intervals



Be prepared to calculate a confidence interval on the Level I exam. Consider a normal distribution with mean μ and standard deviation σ. Each observation has an expected value of μ. If we draw a sample of size *n* from the distribution, the mean of the sample has an expected value of μ. The larger the sample, the closer to μ we expect the sample mean to be. The standard deviation of the means of samples of size *n* is simply $\sigma/\sqrt{n}$ and is called standard error of the sample mean. This allows us to construct a confidence interval for the sample mean for a sample of size *n*.

©2017 Kaplan, Inc.

Example:

Calculate a 95% confidence interval for the mean of a sample of size 25 drawn from a normal distribution with a mean of 8 and a standard deviation of 4.

Answer:

The standard deviation of the means of samples of size 25 is:

$$4/\sqrt{25} = 4/5 = 0.8$$

A 95% confidence interval will extend 1.96 standard deviations above and below the mean, so our 95% confidence interval is:

$8 \pm 1.96 \times 0.8$, 6.432 to 9.568

We believe the mean of a sample of 25 observations will fall within this interval 95% of the time.

With a known variance, the formula for a confidence interval is:

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

In other words, the confidence interval is equal to the mean value, plus or minus the *z*-score that corresponds to the given significance level multiplied by the standard error.

- Confidence intervals and *z*-scores are very important in hypothesis testing, a topic that will be reviewed shortly.

Shortfall Risk and Safety-First Ratio

Shortfall risk. The probability that a portfolio's return or value will be below a specified (target) return or value over a specified period.

Roy's safety-first criterion states that the optimal portfolio minimizes the probability that the return of the portfolio falls below some minimum acceptable "threshold" level.

©2017 Kaplan, Inc.

Roy's safety-first ratio (SFRatio) is similar to the Sharpe ratio. In fact, the Sharpe ratio is a special case of Roy's ratio where the "threshold" level is the risk-free rate of return.

Under both the Sharpe and Roy criteria, the best portfolio is the one that has the largest ratio.

Roy's safety-first ratio can be calculated as:

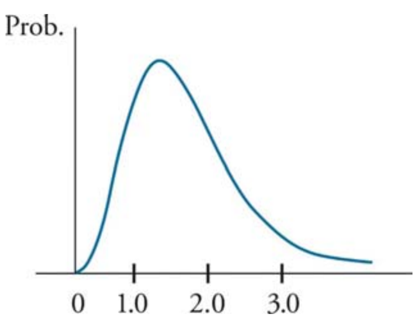
$$\text{SFRatio} = \frac{E\left(R_p\right) - R_L}{\sigma_p}$$

With approximate normality of returns, the SFR is like a *t*-statistic. It shows how many standard deviations the expected return is above the threshold return (R_L). The greater the SFR, the lower the probability that returns will be below the threshold return (i.e., the lower the shortfall risk).

Lognormal Distribution

If x is normally distributed, $Y = e^x$ is lognormally distributed. Values of a lognormal distribution are always positive so it is used to model asset prices (rather than rates of return, which can be negative). The lognormal distribution is positively skewed as shown in the following figure.

Figure 6: Lognormal Distribution



Continuously Compounded Returns

If we increase the number of compounding periods (n) for an annual rate of return, the limit as n goes toward infinity is continuous compounding. For a specific

©2017 Kaplan, Inc.

holding period return (HPR), the relation to the continuously compounded return (CCR) over the holding period is as follows:

$$\text{CCR} = \ln(1+\text{HPR}) = \ln\left(\frac{\text{ending value}}{\text{beginning value}}\right)$$

$$\text{HPR} = \frac{\text{ending value}}{\text{beginning value}} - 1 = e^{\text{CCR}} - 1$$

When the holding period is one year, so that HPR is also the effective annual return, CCR is the annual continuously compounded rate of return.

One property of continuously compounded rates is that they are additive over multiple periods. If the continuously compounded rate of return is 8%, the holding period return over a 2-year horizon is $e^{2(0.08)} - 1$, and \$1,000 will grow to $1{,}000\ e^{2.5(0.08)}$ over two and one-half years.

Simulation

Historical simulation of outcomes (e.g., changes in portfolio values) is done by randomly selecting changes in price or risk factors from actual (historical) past changes in these factors and modeling the effects of these changes on the value of a current portfolio. The results of historical simulation have limitations since future changes may not necessarily be distributed as past changes were.

Monte Carlo simulation is performed by making assumptions about the distributions of prices or risk factors and using a large number of computer-generated random values for the relevant risk factors or prices to generate a distribution of possibly outcomes (e.g., project NPVs, portfolio values). The simulated distributions can only be as accurate as the assumptions about the distributions of and correlations between the input variables assumed in the procedure.

©2017 Kaplan, Inc.

Sampling and Estimation

Cross-Reference to CFA Institute Assigned Reading #11

Know the methods of sampling, sampling biases, and the central limit theorem, which allows us to use sampling statistics to construct confidence intervals around point estimates of population means.

- *Sampling error*: Difference between the sample statistic and its corresponding population parameter:

$$\text{sampling error of the mean} = \bar{x} - \mu$$

- *Simple random sampling*: Method of selecting a sample such that each item or person in the population has the *same likelihood of being included* in the sample.
- *Stratified random sampling*: Separate the population into groups based on one or more characteristics. Take a random sample from each class based on the group size. In constructing bond index portfolios, we may first divide the bonds by maturity, rating, call feature, etc., and then pick bonds from each group of bonds in proportion to the number of index bonds in that group. This insures that our "random" sample has similar maturity, rating, and call characteristics to the index.

Sample Biases

- *Data-mining bias* occurs when research is based on the previously reported empirical evidence of others, rather than on the testable predictions of a well-developed economic theory. Data mining also occurs when analysts repeatedly use the same database to search for patterns or trading rules until one that "works" is found.
- *Sample selection bias* occurs when some data is systematically excluded from the analysis, usually because of the lack of availability.
- *Survivorship bias* is the most common form of sample selection bias. A good example of survivorship bias is given by some studies of mutual fund performance. Most mutual fund databases, like Morningstar's, only include funds currently in existence—the "survivors." Since poorly performing funds are more likely to have ceased to exist because of failure or merger, the survivorship bias in the data set tends to bias average performance upward.
- *Look-ahead bias* occurs when a study tests a relationship using sample data that was not available on the test date.
- *Time-period bias* can result if the time period over which the data is gathered is either too short or too long.

Central Limit Theorem

The *central limit theorem* of statistics states that in selecting simple random samples of size *n* from a *population* with a mean μ and a finite variance σ^2, the sampling distribution of the sample mean approaches a normal probability distribution with mean μ and a variance equal to σ^2/n as the sample size becomes large.

The central limit theorem is extremely useful because the normal distribution is relatively easy to apply to hypothesis testing and to the construction of confidence intervals.

Specific inferences about the population mean can be made from the sample mean, *regardless of the population's distribution*, as long as the sample size is sufficiently large.

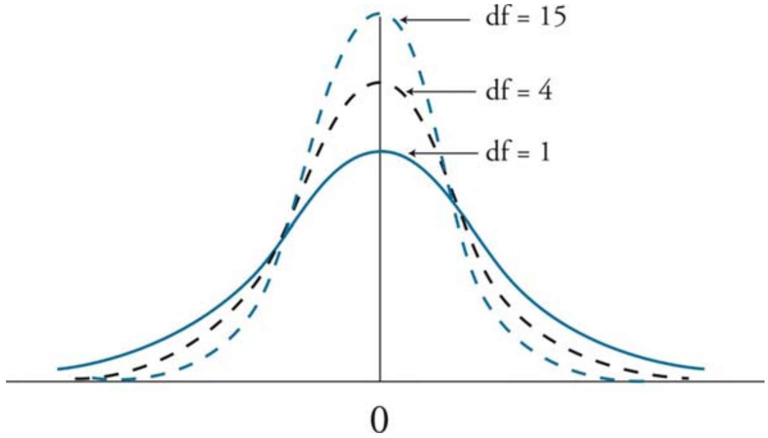
Student's *t*-Distribution

- Symmetrical (bell shaped).
- Defined by single parameter, degrees of freedom (df), where df = n – 1 for hypothesis tests and confidence intervals involving a sample mean.
- Has fatter tails than a normal distribution; the lower the df, the fatter the tails and the wider the confidence interval around the sample mean for a given probability that the interval contains the true mean.
- As sample size (degrees of freedom) increases, the *t*-distribution approaches normal distribution.

Student's t-distribution is similar in concept to the normal distribution in that it is bell-shaped and symmetrical about its mean. The *t-distribution* is appropriate when working with small samples (n < 30) from populations with *unknown variance* and normal, or approximately normal, distributions. It may also be appropriate to use the *t*-distribution when the population variance is unknown and the sample size is large enough that the central limit theorem will assure the sampling distribution is approximately normal.

©2017 Kaplan, Inc.

Figure 7: Student's *t*-Distribution and Degrees of Freedom



For questions on the exam, make sure you are working with the correct distribution. You should memorize the following table:

Figure 8: Criteria for Selecting Test Statistic

When sampling from a:	*Test Statistic*	
	Small Sample *(n < 30)*	*Large Sample* *(n ≥ 30)*
Normal distribution with *known* variance	z-statistic	z-statistic
Normal distribution with *unknown* variance	t-statistic	t-statistic*
Nonnormal distribution with *known* variance	not available	z-statistic
Nonnormal distribution with *unknown* variance	not available	t-statistic**

* The z-statistic is the standard normal, ±1 for 68% confidence, et cetera.
** The z-statistic is theoretically acceptable here, but use of the t-statistic is more conservative.

Hypothesis Testing

Cross-Reference to CFA Institute Assigned Reading #12

Hypothesis. Statement about a population parameter that is to be tested. For example, "The mean return on the S&P 500 Index is equal to zero."

Steps in Hypothesis Testing

- State the hypothesis.
- Select a test statistic.

©2017 Kaplan, Inc.

- Specify the level of significance.
- State the decision rule for the hypothesis.
- Collect the sample and calculate statistics.
- Make a decision about the hypothesis.
- Make a decision based on the test results.

Null and Alternative Hypotheses

The *null hypothesis*, designated as H_0, is the hypothesis the researcher wants to reject. It is the hypothesis that is actually tested and is the basis for the selection of the test statistics. Thus, if you believe (seek to show that) the mean return on the S&P 500 Index is different from zero, the null hypothesis will be that the mean return on the index *equals* zero.

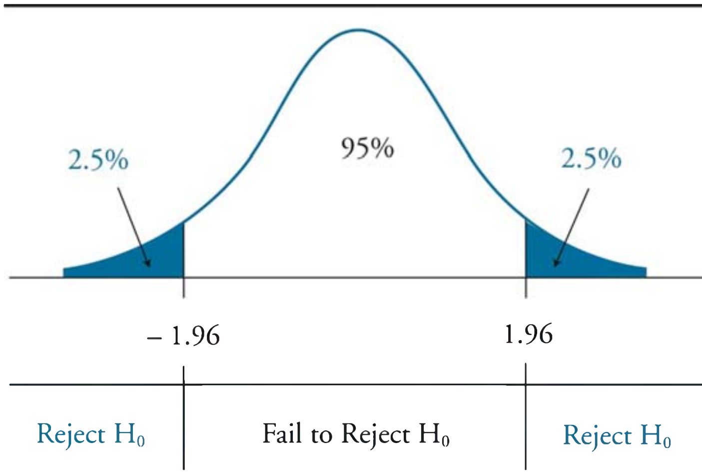
The *alternative hypothesis*, designated H_a, is what is concluded if there is sufficient evidence to reject the null hypothesis. It is usually the alternative hypothesis you are really trying to support. Why? Since you can never really prove anything with statistics, when the null hypothesis is rejected, the implication is that the (mutually exclusive) alternative hypothesis is valid.

Two-Tailed and One-Tailed Tests

Two-tailed test. Use this type of test when testing a parameter to see if it is different from a specified value:

$$H_0: \mu = 0 \text{ versus } H_a: \mu \neq 0$$

Figure 9: Two-Tailed Test: Significance = 5%, Confidence = 95%



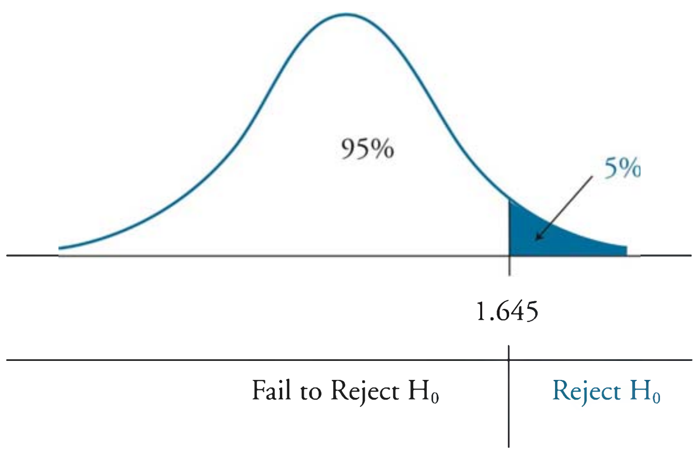
©2017 Kaplan, Inc.

One-tailed test. Use this type of test when testing a parameter to see if it is *above* or *below* a specified value:

H_0: $\mu \leq 0$ versus H_a: $\mu > 0$, or
H_0: $\mu \geq 0$ versus H_a: $\mu < 0$

With respect to the first hypothesis, $\mu \leq 0$, we will reject it only if the test statistic is significantly greater than zero (in the right-hand tail of the distribution). Thus, we call it a one-tailed test.

Figure 10: One-Tailed Test: Significance = 5%, Confidence = 95%



Test Statistic

A *test statistic* is calculated from sample data and is compared to a critical value to evaluate H_0. The most common test statistics are the *z*-statistic and the *t*-statistic. Which statistic you use to perform a hypothesis test will depend on the properties of the population and the sample size as noted above.

- Critical values come from tables and are based on the researcher's desired level of significance. As the level of significance (the α) gets smaller, the critical value gets larger and it becomes more difficult to reject the null hypothesis.
- If the test statistic exceeds the critical value (or is outside the range of critical values), the researcher rejects H_0.

Type I and Type II Errors

When testing a hypothesis, there are two possible types of errors:

- *Type I error*. Rejection of the null hypothesis when it is actually true.
- *Type II error*. Failure to reject the null hypothesis when it is actually false.

©2017 Kaplan, Inc.

The *power of a test* is 1 – P(Type II error). The more likely that a test will reject a false null, the more powerful the test. A test that is unlikely to reject a false null hypothesis has little power.

Significance Level (α)

The *significance level* is the probability of making a Type I error (rejecting the null when it is true) and is designated by the Greek letter alpha (α). You can think of this as the probability that the test statistic will exceed or fall below the critical values by chance even though the null hypothesis is true. A significance level of 5% (a = 0.05) means there is a 5% chance of rejecting a true null hypothesis.

Figure 11: Errors in Hypothesis Testing

Type I and Type II Errors in Hypothesis Testing		
Decision	True Condition	
	H_0 is true	H_0 is false
Do not reject H_0	Correct decision	Incorrect decision Type II error
Reject H_0	Incorrect decision Type I error Significance level, α, = P(Type I error)	Correct decision Power of the test = 1 – P(Type II error)

Economically Meaningful Results

A test may indicate a significant statistical relationship (a statistically meaningful result) which is not economically significant. This is often the case when the gains from exploiting the statistical relation are small in an absolute sense so that the costs of a strategy to exploit the relation are greater than the expected gains from the strategy.

Other Hypothesis Tests

A test of the equality of the means of two independent normally distributed populations is a *t*-test based on the difference in sample means divided by a standard deviation which is calculated in one of two ways, depending on whether the variances of the two populations are assumed to be equal or not.

When random variables from two populations are dependent, the appropriate test is a *mean differences* or *paired comparisons* test. The test statistic is a *t*-statistic based

 ©2017 Kaplan, Inc.

on the average (mean) of the differences in the sample of the paired values of the two random variables, divided by the standard deviation of the differences between the sample pairs.

A test of whether the population variance of a normal distribution is equal to a specific value is based on the ratio of the sample variance to the hypothesized variance. The test statistic follows a Chi-square distribution and is a two-tailed test.

A test of whether the variances of two normal populations are equal is based on the ratio of the larger sample variance to the smaller sample variance. The appropriate test is an *F*-test (two-tailed), but by putting the larger sample variance in the numerator, values of the test statistic below the lower critical value are ruled out, and only the upper critical value of the *F*-statistic need be considered.

Figure 12 summarizes the test statistics used for each type of hypothesis test.

Figure 12: Types of Test Statistics

Hypothesis tests of:	*Use a:*
One population mean	t-statistic or *Z*-statistic
Two population means	t-statistic
One population variance	Chi-square statistic
Two population variances	F-statistic

Parametric and Nonparametric Tests

Parametric tests, like the *t*-test, *F*-test, and chi-square test, make assumptions regarding the distribution of the population from which samples are drawn.

Nonparametric tests either do not consider a particular population parameter or have few assumptions about the sampled population. Runs tests (which examine the pattern of successive increases or decreases in a random variable) and rank correlation tests (which examine the relation between a random variable's relative numerical rank over successive periods) are examples of nonparametric tests.

Technical Analysis

Cross-Reference to CFA Institute Assigned Reading #13

This topic review presents many different technical analysis tools. Don't try to memorize them all. Focus on the basics of technical analysis and its underlying assumptions.

©2017 Kaplan, Inc.

Assumptions of Technical Analysis

- Values, and thus prices, are determined by supply and demand.
- Supply and demand are driven by both rational and irrational behavior.
- Price and volume reflect the collective behavior of buyers and sellers.
- While the causes of changes in supply and demand are difficult to determine, the actual shifts in supply and demand can be observed in market price behavior.

Advantages of Technical Analysis

- Based on observable data (price and volume) that are not based on accounting assumptions or restatements.
- Can be used for assets that do not produce cash flows, such as commodities.
- May be more useful than fundamental analysis when financial statements contain errors or are fraudulent.

Disadvantages of Technical Analysis

- Less useful for markets that are subject to outside intervention, such as currency markets, and for markets that are illiquid.
- Short covering can create positive technical patterns for stocks of bankrupt companies.
- Cannot produce positive risk-adjusted returns over time when markets are weak-form efficient.

Types of Charts

Except for point and figure charts, all of the following chart types plot price or volume on the vertical axis and time (divided into trading periods) on the horizontal axis. Trading periods can be daily, intraday (e.g., hourly), or longer term (e.g., weekly or monthly).

Line chart: Closing prices for each trading period are connected by a line.

Bar chart: Vertical lines from the high to the low price for each trading period. A mark on the left side of the line indicates the opening price and a mark on the right side of the vertical line indicates the closing price.

Candlestick chart: Bar chart that draws a box from the opening price to the closing price on the vertical line for each trading period. The box is empty if the close is higher than the open and filled if the close is lower than the open.

©2017 Kaplan, Inc.

Volume chart: Vertical line from zero to the number of shares (bonds, contracts) exchanged during each trading period. Often displayed below a bar or candlestick chart of the same asset over the same range of time.

Point and figure chart: Displays price trends on a grid. Price is on the vertical axis, and each unit on the horizontal axis represents a change in the direction of the price trend.

Relative strength chart: Line chart of the ratios of closing prices to a benchmark index. These charts illustrate how one asset or market is performing relative to another. Relative strength charts are useful for performing intermarket analysis and for identifying attractive asset classes and assets within each class that are outperforming others.

Trend, Support, and Resistance

A market is in an uptrend if prices are consistently reaching higher highs and retracing to higher lows. An uptrend indicates demand is increasing relative to supply. An upward sloping trendline can be drawn that connects the low points for a stock in an uptrend.

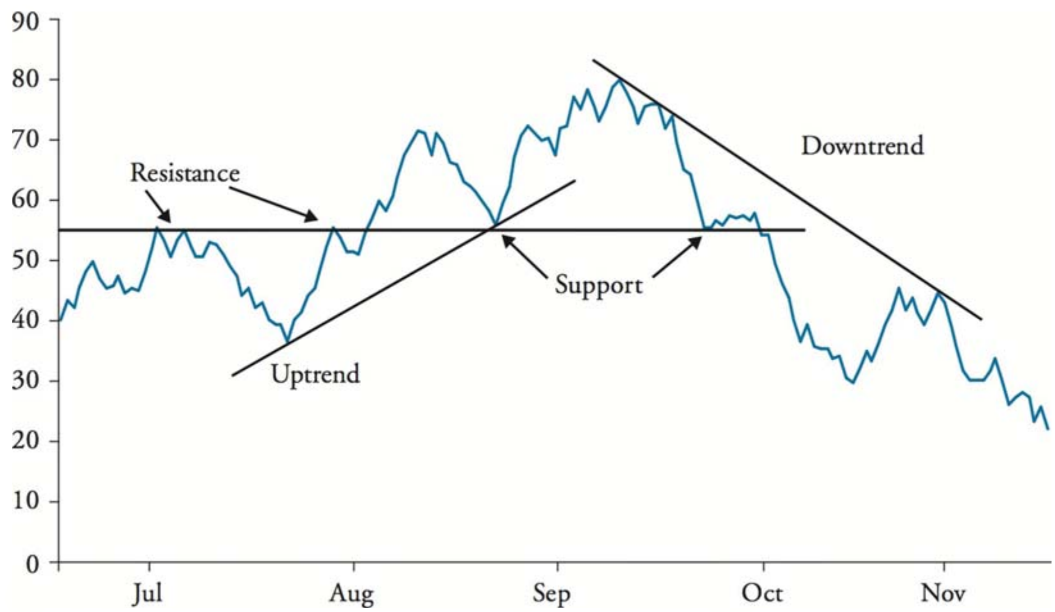
A market is in a downtrend if prices are consistently reaching lower lows and retracing to lower highs. A downtrend means supply is increasing relative to demand. A downward sloping trendline can be drawn that connects the high points in a downtrend.

Support and resistance levels are prices at which technical analysts expect supply and demand to equalize. Past highs are viewed as resistance levels, and past lows are viewed as support levels. Trendlines are also thought to indicate support and resistance levels.

The *change in polarity principle* is based on a belief that breached support levels become resistance levels, and breached resistance levels become support levels.

©2017 Kaplan, Inc.

Figure 13: Trendlines, Support, and Resistance



Common Chart Patterns

Reversal patterns: Head-and-shoulders; double top; triple top; inverse head-and-shoulders; double bottom; triple bottom. These price patterns are thought to indicate that the preceding trend has run its course and a new trend in the opposite direction is likely to emerge.

Continuation patterns: Triangles; rectangles; flags; pennants. These indicate temporary pauses in a trend which is expected to continue (in the same direction).

Technical analysts often use the sizes of both of these types of patterns to estimate subsequent target prices for the next move.

Price-based Indicators

Moving average lines are a frequently used method to smooth the fluctuations in a price chart. A 20-day moving average is the arithmetic mean of the last 20 closing prices. The larger number of periods chosen, the smoother the resulting moving average line will be. Moving average lines can help illustrate trends by smoothing short-term fluctuations, but when the number of periods is large, a moving average line can obscure changes in trend.

Bollinger bands are drawn a given number of standard deviations above and below a moving average line. Prices are believed to have a higher probability of falling (rising) when they are near the upper (lower) band.

©2017 Kaplan, Inc.

Momentum oscillators include the rate of change oscillator, the Relative Strength Index (RSI), moving average convergence/divergence (MACD) lines, and stochastic oscillators.

Technical analysts use price-based indicators to identify market conditions that are overbought (prices have increased too rapidly and are likely to decrease in the near term) or oversold (prices have decreased too rapidly and are likely to increase in the near term). They also use charts of momentum oscillators to identify convergence or divergence with price trends. Convergence occurs when the oscillator shows the same pattern as prices (e.g., both reaching higher highs). Divergence occurs when the oscillator shows a different pattern than prices (e.g., failing to reach a higher high when the price does). Convergence suggests the price trend is likely to continue, while divergence indicates a potential change in trend in the near term.

Sentiment and Flow of Funds Indicators

Technical analysts also use indicators based on investors' bullish (investors expect prices to increase) or bearish (investors expect prices to decrease) sentiment. Some technical analysts interpret these indicators from a contrarian perspective. Contrarians believe markets get overbought or oversold because most investors tend to buy and sell at the wrong times, and thus it can be profitable to trade in the opposite direction from current sentiment.

Sentiment indicators include the following:

- *Put/call ratio:* Put option volume divided by call option volume.
- *Volatility index* (VIX): Measure of volatility on S&P 500 stock index options.
- *Short interest ratio:* Shares sold short divided by average daily trading volume.
- Amount of *margin debt* outstanding.
- *Opinion polls* that attempt to measure investor sentiment directly.

High levels of the put/call ratio, VIX, and short interest ratio indicate bearish market sentiment, which contrarians interpret as bullish. High levels of margin debt indicate bullish sentiment, which contrarians interpret as bearish.

Indicators of the flow of funds in the financial markets can be useful for identifying changes in the supply and demand for securities. These include the Arms index or short-term trading index (TRIN), which measures funds flowing into advancing and declining stocks; margin debt (also used as a sentiment indicator); new and secondary equity offerings; and mutual fund cash as a percentage of net assets.

©2017 Kaplan, Inc.

Cycles and Elliott Wave Theory

Some technical analysts apply cycle theory to financial markets in an attempt to identify cycles in prices. Cycle periods favored by technical analysts include 4-year presidential cycles related to election years in the United States, decennial patterns or 10-year cycles, 18-year cycles, and 54-year cycles called Kondratieff waves.

One of the more developed cycle theories is the Elliott wave theory which is based on an interconnected set of cycles that range from a few minutes to centuries. According to Elliott wave theory, in an uptrend the upward moves in prices consist of five waves and the downward moves occur in three waves. If the prevailing trend is down, the downward moves have five waves and the upward moves have three waves. Each of these waves is composed of smaller waves that exhibit the same pattern.

The sizes of these waves are thought to correspond with ratios of Fibonacci numbers. Fibonacci numbers are found by starting with 0 and 1, then adding each of the previous two numbers to produce the next (0, 1, 1, 2, 3, 5, 8, 13, 21, and so on). Ratios of consecutive Fibonacci numbers converge to 0.618 and 1.618 as the numbers in the sequence get larger.

 ©2017 Kaplan, Inc.