

# FN241: Session 2-3

## Part I: Probability and Statistical Concepts

Winai Homsombat

Bachelor of Economics, International Program

Thammasat University

# Statistical Basic Concepts

# Populations and Samples

- The subset of data used in statistical inference is known as a **sample** and the larger body of data is known as the **population**.
  - The **population** is defined as all members of the group in which we are interested.

# Parameters and Sample Statistics

A population has parameters, and a sample has statistics.

- Descriptive statistics that characterize population values are called **parameters**.
  - Examples: mean, median, mode, variance, skewness, kurtosis
- Descriptive statistics that characterize samples are known as **sample statistics**.
  - Examples: sample mean, sample median, sample variance
- By convention, we often omit the term “sample” in front of sample statistics, a practice that can lead to confusion when discussing both the sample and the population.

# Measures of central tendency

These measures describe where the data are centered.

- Arithmetic Mean
  - The **arithmetic mean** is the sum of the observations divided by the number of observations.
  - Population mean →
  - Sample mean →
  - The sample mean is often interpreted as the fulcrum, or center of gravity, for a given set of data.
  - **Cross-sectional data** occur across different observation types at one point in time, and **time-series data** occur for the same unit of observation across time.

# Measures of central tendency

## Example: Cross-Sectional Sample Mean Return

Country	Return	Country	Return
Austria	-2.97%	Italy	-23.64%
Belgium	-29.71%	Netherlands	-34.27%
Denmark	-29.67%	Norway	-29.73%
Finland	-41.65%	Portugal	-28.29%
France	-33.99%	Spain	-29.47%
Germany	-44.05%	Sweden	-43.07%
Greece	-39.06%	Switzerland	-25.84%
Ireland	-38.97%	United Kingdom	-25.66%

Source: CFA Institution

# Measures of central tendency

These measures also describe where the data are centered.

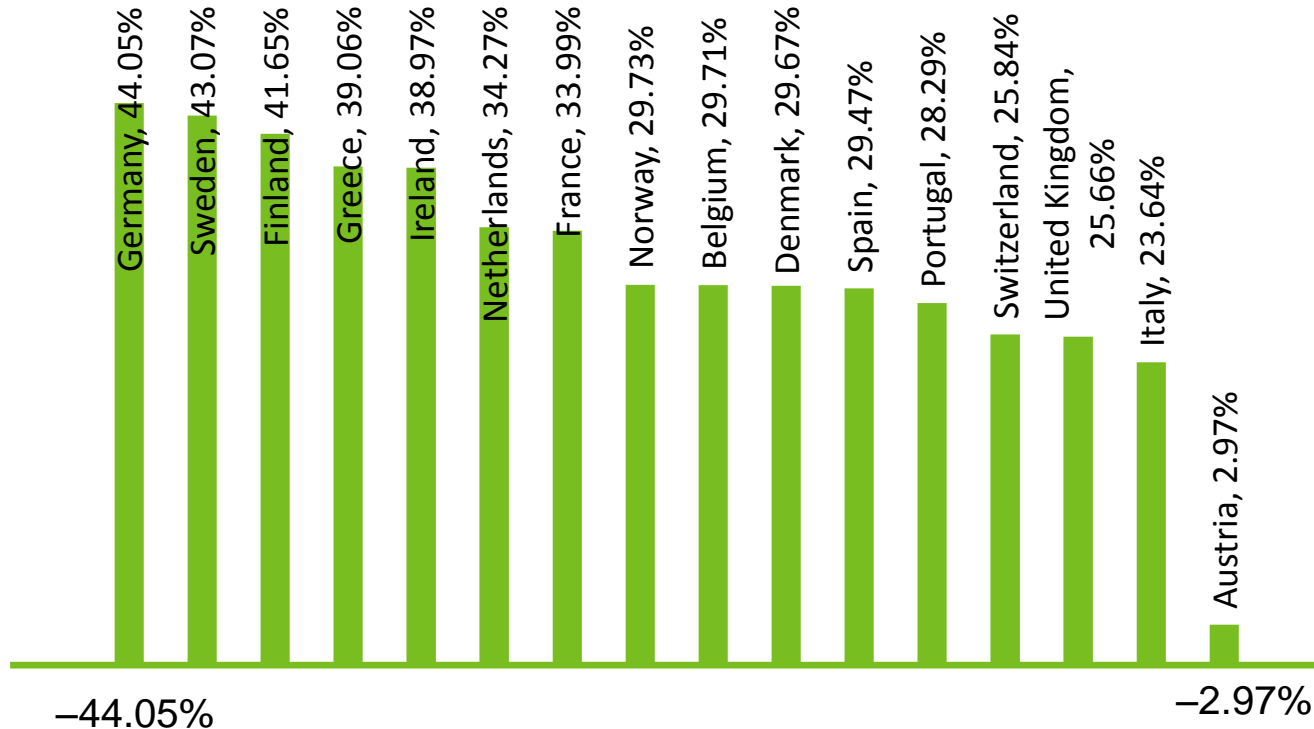
- **Weighted Mean:**

- The sum of the observations times each observation's weight (proportional representation in the sample), where the weight is chosen to meet a statistical or financial goal. Example: Portfolio return

- **Geometric Mean:**

# Measures of central tendency

Mean as a center of gravity for the data object



# Measures of dispersion

Dispersion measures variability around a measure of central tendency. If mean return represents reward, then dispersion represents risk.

- Variance is the average squared deviation from the mean.
  - Population variance:
  - Sample variance:
- Sample variance is “penalized” by dividing by  $n - 1$  instead of  $n$  to account for the fact that the measure of central tendency used,  $\bar{X}$ , is an estimate of the true population parameter,  $\mu$ , and so has some uncertainty associated with it.
- Standard deviation is the square root of variance.

# Measures of dispersion

## Focus on: Sample Standard Deviation

Country	Return	Squared Deviation from Mean
Germany	-44.05%	0.016384
Sweden	-43.07%	0.013971
Finland	-41.65%	0.010816
Greece	-39.06%	0.00610
...	...	...
Austria	-2.97%	0.0780

# Semivariance

We are often concerned with measures of risk that focus on the “downside” of the possible outcomes - in other words, the losses.

- Semivariance is the average squared deviation below the mean.
  - Semideviation is the square root of semivariance.
  - Both are a measure of dispersion focusing only on those observations below the mean.
  
- Target semivariance, by analogy, is the average squared deviation below some specified target rate,  $B$ , and represents the “downside” risk of being below the target,  $B$ .

# Chebyshev's Inequality

- This expression gives the minimum proportion of values,  $p$ , within  $k$  standard deviations of the mean for any distribution whenever  $k > 1$ .

$k$	Interval around the Mean	$p$
1.25	$\bar{X} \pm 1.25s$	0.36
1.50	$\bar{X} \pm 1.50s$	0.56
2.00	$\bar{X} \pm 2.00s$	0.75
2.50	$\bar{X} \pm 2.50s$	0.84
3.00	$\bar{X} \pm 3.00s$	0.89
4.00	$\bar{X} \pm 4.00s$	0.94

# Chebyshev's Inequality

- Focus on: Calculating Proportions Using Chebyshev's Inequality
- For our country data, the mean is  $-31.25\%$  and the sample standard deviation is  $9.95\%$ .

Lower cutoff at 1.25 standard deviations:

Upper cutoff at 1.25 standard deviations:

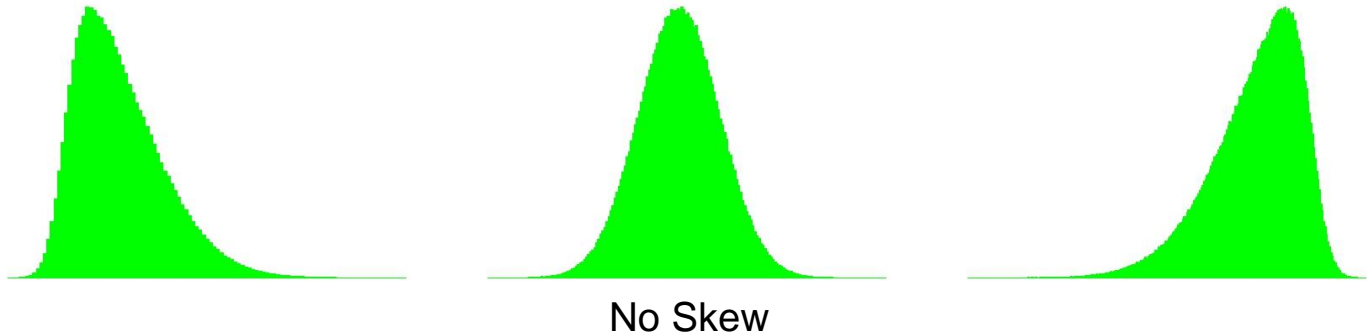
...

...

$k$	Lower Cutoff	Upper Cutoff	Actual $p$	Chebyshev's $p$
1.25	$-43.69\%$	$-18.81\%$	0.875	0.36
1.50	$-46.18\%$	$-16.32\%$	0.938	0.56
2.00	$-51.16\%$	$-11.34\%$	0.938	0.75
2.50	$-56.13\%$	$-6.37\%$	0.938	0.84
3.00	$-61.11\%$	$-1.39\%$	0.938	0.89
4.00	$-71.07\%$	$8.57\%$	1.000	0.94

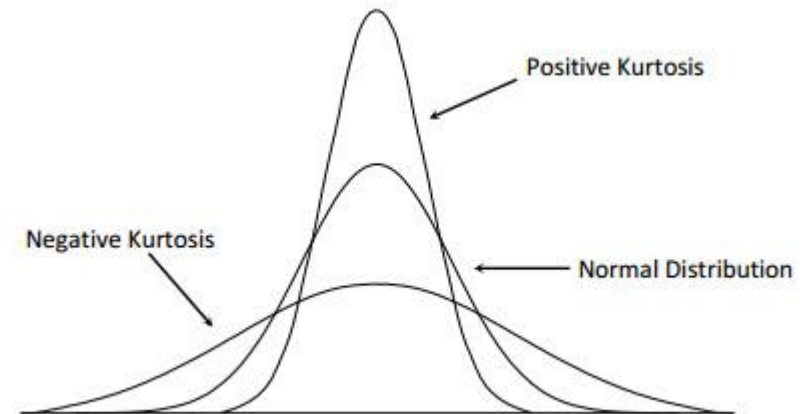
# Skewness:

- The degree of symmetry in the dispersion of values around the mean is known as skewness.
- If observations are equally dispersed around the mean, the distribution is said to be symmetrical.
- If the distribution has a long tail on one side and a “fatter” distribution on the other side, it is said to be skewed in the direction of the long tail.



# Kurtosis:

- Kurtosis measures the relative amount of “peakedness” as compared with the normal distribution, which has a kurtosis of 3.
  - We typically express this measure in terms of excess kurtosis being the observed kurtosis minus 3.
  - Distributions are referred to as being
    1. Leptokurtic  
(more peaked than the normal; fatter tails)
    2. Platykurtic  
(less peaked than the normal; thinner tails) or
    3. Mesokurtic (equivalent to the normal).



# Probability Concepts

# Fundamental Concepts

- A variable is **random** if its outcome is uncertain, where an **outcome** is an observable future value of the variable.
- An **event** is the specified set of possible outcomes of a random variable.
  - Events are **mutually exclusive** when the possible future outcomes can only occur one at a time and **exhaustive** when the set of outcomes includes every possible value the variable could take in the future.
    - **Example:** The future size of a dividend can be stated as a mutually exclusive and exhaustive event wherein dividends increase, decrease, or remain unchanged.
  - When the occurrence of one event does not affect the probability of the occurrence of another event, we say the events are **independent**.
  - Events that are not independent are **dependent**.

# Probability

- Probability is the fundamental building block of statistics.
- Probability is a number between 0 and 1 that describes **the chance that a stated event from the set of possible outcomes** will occur.
- A probability distribution is the set of probabilities and their associated outcomes that describes **all possible outcomes and their associated probabilities**.
- We typically use  $P(E)$  to denote the probability of event  $E$ .
  - Properties of probability
    1. All probabilities must lie between 0 and 1:
    2. For  $n$  mutually exclusive and exhaustive events, the sum of all probabilities must equal 1:

# Types of Probability

## Sources of probabilities

- In practice, we observe a number of different types of probability.
  - A **subjective probability** is a personal assessment of the likelihood of an event or set of events occurring in the future and is so named because it relies on the subjective judgment of the person making the assessment.
  - An **empirical probability** is one that is estimated from observed data, typically using the relative frequency at which an event or set of events has occurred in the past.
  - An **a priori probability** is one whose values are obtained from mathematical or logical analysis.

# Probability as odds

Probabilities are often stated as odds for or against a given event occurring.

Odds for:

Odds against:

Tire industry analysts are interested in the odds of a removal of import restrictions on Chinese tires. Your analyst has assessed the probability of removal as 60%. What are the odds for and against removal?

Odds for:

Odds against:

# Joint Probability

- When two events occur, the combined probability of their occurrence is known as the joint probability.
  
- Notation:  $P(AB)$
- Language: the probability of  $A$  and  $B$
- The calculation of a joint probability is governed by a set of probability rules, as are the calculations of the probability of other combinations of events.

# Working with probabilities

- We use the **multiplication rule** to assess the joint probability of  $A$  and  $B$  occurring.
  - Note that from this equation, we can calculate the conditional probability of  $A$  given  $B$ , as long as the probability of  $B$  isn't zero.
- We use the **addition rule** to assess the probability that  $A$  or  $B$  or both occur:

# Multiplication rule

**Example:** Recall that our probability of relaxed trade restrictions has been estimated at 60%.

If the probability of reduced sales, given that the trade restrictions are relaxed, is 80%, then the probability of relaxed trade restrictions and reduced sales is

# Addition rule

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

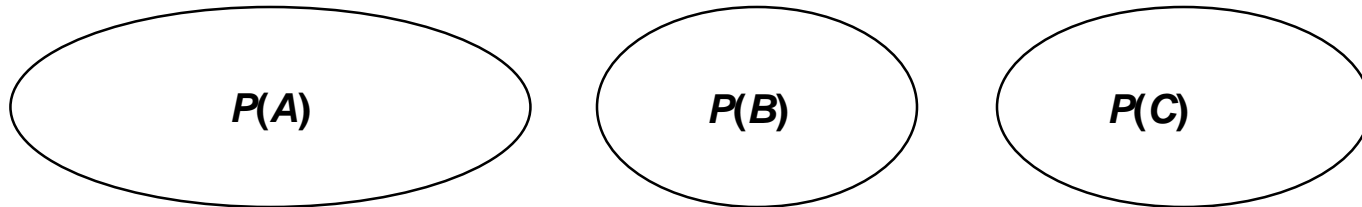
**Example:** If the probability of relaxed import restrictions is 0.60 and the probability of a trade war is 0.10, then the probability of relaxed trade restrictions or a trade war is 0.65 when the joint probability of a trade war and relaxed trade restrictions is 0.05.

# Dependent and Independent events

- When events are independent, the occurrence of one does not affect the probability of the other. In other words,  $P(A|B) = P(A)$  and  $P(B|A) = P(B)$ .
- To determine the joint probability, we use the multiplication rule:

$$P(AB) = P(A)P(B)$$

or, more generally,



# Joint probability of independent events

## Example:

- Recall our analyst's estimated probability of relaxed import restrictions (60%).
- If  $A$  is the probability of a heavy winter snowfall (40%) and  
 $B$  is the independent probability of relaxed restrictions:
- Then the joint probability of relaxed import restrictions and a heavy winter snowfall is:

# Total probability rule

- Recall that we have a 60% (40%) chance of trade restrictions being relaxed (maintained).
  - $P(RR) =$  , where  $RR$  is relaxed restrictions.
  - $P(NR) =$  , where  $NR$  is no relaxation.
- Our analysts believe that the stock price will decrease if trade restrictions are relaxed with a probability of 30% and that they will decrease if trade restrictions are not relaxed with a probability of 15%.
  - $P(DS|RR) =$
  - $P(DS|NR) =$
- What is the probability of a decrease in stock prices?

# Expected value

## Random Variable

- The expected value of a random variable is the probability-weighted average of the possible outcomes for that variable.

**Example:** We anticipate that there is a 15% chance that next year's return on holding Cleveland Corp will be 4%, a 60% chance it will be 6%, and a 25% chance it will be 8%. What is the expected return on Cleveland Corp stock?

# Variance

- The variance of a random value is the sum of the squared deviations from the expected value weighted by their associated probabilities.
  - This value is a measure of the dispersion of possible values.
  - Because it has units that are squared, it is not easy to interpret. Accordingly, we use its positive square root, standard deviation, more often because it also measures dispersion but has the same units as expected value.
- The standard deviation of returns for Cleveland Corp. is then:

# Conditional Expectations

- The total probability rule applies to expected values just as it does any mutually exclusive and exhaustive set of possible outcomes across a set of states.
  
- This equation allows us to calculate the expected value of a random variable ( $X$ ) as a function of the probabilities of future possible states,  $P(S)$ , and the conditional value of the expected value of  $X$  in those states,  $E(X|S)$ .

# Conditional Expected value

## Example:

Example: Recall that we have a 60% chance of relaxed trade restrictions and, therefore, a 40% of maintaining them. If we expect the Cleveland Corp. stock to return 6% if trade restrictions are maintained and lose 11% if they are relaxed, what is the expected change in return for Cleveland Corp.?

# Tree diagram

A visual representation of the future possible outcomes and associated probabilities of a random variable.

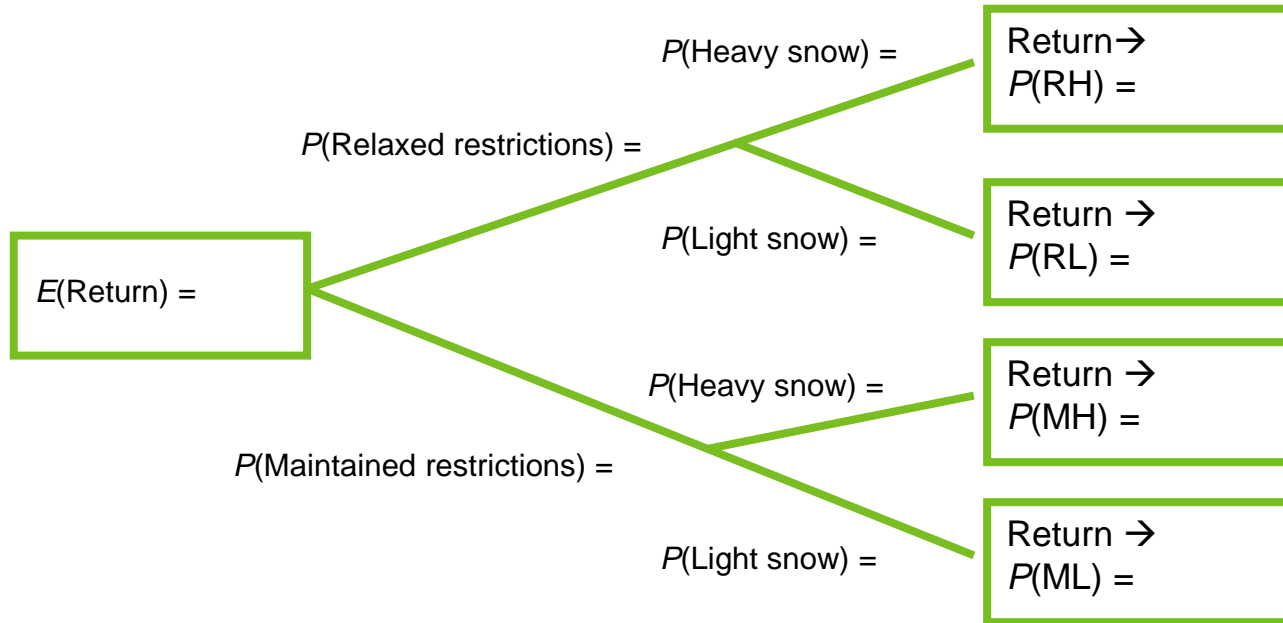
- Recall that we have a 60% (40%) chance of trade restrictions being relaxed (maintained) and a 40% (60%) chance of a heavy (light) winter snowfall.
- Our analyst projects that if trade restrictions are relaxed and there is a light snowfall, Cleveland Corp. will have a stock return of 15%, but if restrictions are relaxed and there is a heavy snowfall, there will be a stock return of 5%. If trade restrictions are not relaxed and there is a light snowfall, we will have a stock return of 2%, but if we have a heavy snowfall, we will have stock return of 12%.

How do we represent the possible impact of snowfall and trade restrictions on the return of Cleveland Corp. stock?

What is our overall expected stock return on Cleveland Corp.?

# Tree diagram

Example:



# Covariance

Covariance and correlation are both measures of the extent to which two random variables move together.

- Covariance is the expected value of the product of each variable's deviation from its respective mean.

$$\sigma_{X,Y} = E\{[X - E(X)][Y - E(Y)]\}$$

$$\sigma_{R_i,R_j} = \sum_{i=1}^n P(R_i)[R_i - E(R_i)][R_j - E(R_j)]$$

# Correlation

Covariance and correlation are both measures of the extent to which two random variables move together.

- Correlation is a scaled transformation of covariance wherein the extent of comovement is measured along a scale from exactly the same movement in the same direction to exactly the same movement in opposite directions.
  - When two variables move the same degree in opposing directions, they are said to be perfectly negatively correlated.
  - When two variables move the same degree in the same direction, they are said to be perfectly positively correlated.
  - When there is absolutely no commonality of movement, the variables are said to be uncorrelated.

# Covariance and Correlation

## Example:

- We anticipate that there is a 15% chance that next year's stock returns for Cleveland Corp. will be 4%, a 60% chance they will be 6%, and a 25% chance they will be 8%. The standard deviation of returns is 1.249%, and the expected value is 6.2%.
- We anticipate that the same probabilities and states are associated with a 2% return for High Noon Inc., a 3% return, and a 3.5% return. The standard deviation of High Noon Inc. returns is then 0.46%, and its expected value is 2.975%.
  - What is the covariance between Cleveland and High Noon returns?
  - What is the correlation between Cleveland and High Noon returns?

# Counting Problems

The foundation for many probability statements lies in enumerating the possible outcomes in order to assess the likelihood of their occurrence.

- If a task can be done in  $n_1$  ways, a second task can be done in  $n_2$  ways, a third task in  $n_3$  ways, and so on, we can enumerate the number of ways in which the tasks can be done by taking the product of all the ways all of the tasks can be done. If there are  $k$  tasks, this mean the number of ways the tasks can be done is
- **Example:** You have asked your analyst to perform four different analyses on the firm he covers. The first process covers six steps; the second, four; the third, three; and the last, seven. How many possible ways can the analyst perform the steps?

# Factorial

The mathematical factorial function is used to enumerate the number of ways in which every member of a group can be assigned a single slot when the number of slots equals the number of members in the group.

- In short, once the first ranking is made, there are  $n - 1$  firms left to be ranked; after the second,  $n - 2$ ; and so on, until all firms are ranked.
- The symbol “!” is used to denote the factorial operation.
  - **Example:** Your firm covers seven firms, and you have asked your analyst to rank the seven firms. How many different possible combinations of rankings are possible for the seven firms?

# Combinatorial

The mathematical process by which every member of a group can be assigned to one of two slots when there are only two categories is called a “combinatorial process.”

- The formula for this calculation is called the “binomial formula” or “combination formula.”
- It is often described as the way we can choose  $r$  objects from  $n$  objects when the order of the  $r$  objects does not matter.
  - **Example:** You have asked the analyst to rate each of the 10 firms he covers as a “watchlist” or “not-watchlist” firm but indicated there can only be 4 watchlist firms. How many possible combinations are there of the resulting groupings?
- The notation in parenthesis at the beginning of the expression is read as “ $n$  choose  $r$ .”

# Permutations

The mathematical process by which every member of a group can be assigned to a set of slots when the order of the slots does matter is known as a permutation.

- **Example:** You have asked the analyst to select and rank four of the six firms he left off the watchlist. How many possible combinations of rankings are there of the top four non-watchlist firms?

# Counting Problems

- **Example:** You are assigned the task of selecting 6 “interesting” firms from a set of 24 firms your boss has screened. How many different potential sets of six firms can you select?
  
  
  
  
  
  
  
  
  
  
- **Example:** Your boss has changed the task, asking you to rank order the 6 most interesting firms out of the 24 selected. How many different possible rankings can you select?

# How Do We Choose Which Counting Method to Use?

1. Does the task that I want to measure have a finite number of possible outcomes? If the answer is yes, you may be able to use a tool in this section, and you can go to the second question. If the answer is no, the number of outcomes is infinite, and the tools in this section do not apply.
2. Do I want to assign every member of a group of size  $n$  to one of  $n$  slots (or tasks)? If the answer is yes, use  $n$  factorial. If the answer is no, go to the third question.
3. Do I want to count the number of ways to apply one of three or more labels to each member of a group? If the answer is yes, use the multinomial formula. If the answer is no, go to the fourth question.
4. Do I want to count the number of ways that I can choose  $r$  objects from a total of  $n$ , when the order in which I list the  $r$  objects does not matter (can I give the  $r$  objects a label)? If the answer to these questions is yes, the combination formula applies. If the answer is no, go to the fifth question.
5. Do I want to count the number of ways I can choose  $r$  objects from a total of  $n$ , when the order in which I list the  $r$  objects is important? If the answer is yes, the permutation formula applies. If the answer is no, go to question 6.
6. Can the multiplication rule of counting be used? If it cannot, you may have to count the possibilities one by one, or use more advanced techniques than those presented here.

# Summary

- The concepts of probability and random variables lie at the center of inferential and descriptive statistics, and thus form an important basis for the chapters to come.
  - In the descriptive statistics area, measures of expected return, variance, and covariance play a central role in portfolio applications in investments.
- We can use a variety of probability tools and rules to determine the likelihood of a given event or set of events in the future.
- We can use counting rules to determine the probability of a given event or set of events in combination by enumerating the possible potential outcomes.

Question?