CBK Applied Decision Methods Review

Introduction by Dr. Robert G. Brown, PhD, PMP

I’ve been a business owner, a sub-chapter S business stockholder, and a director for a multi-national firm. I’ve been managing businesses for over 35 years. In my opinion, the areas emphasized by the Educational Testing Service (ETS) are spot on for the quantitative skills that you’ll need to make more money and keep more money for yourself and for your organization. This review document is organized around the subjects you’ll see on the ETS Business Test Copyright © 2014 by Educational Testing Service. The Applied Decision Methods portion of the test comprises about 11% of the total ETS MFT score.

Part A Probability and Statistics

1. Set Operations
   This is not a standalone document. It is coordinated with your Quantitative Analysis for Management text. Render, Stair, & Hanna (2012, pp. 26-28) note:

   Venn Diagrams

   - Events that are mutually exclusive
     \[ P(A \text{ or } B) = P(A) + P(B) \]
   - Events that are not mutually exclusive
     \[ P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \]

2. Three Types of Probabilities
   Marginal (or simple) probability is just the probability of a single event occurring.
   \[ P(A) \]

   Joint probability is the probability of two or more events occurring and is equal to the product of their marginal probabilities for independent events.
   \[ P(AB) = P(A) \times P(B) \]

  Conditional probability is the probability of event B given that event A has occurred.
   \[ P(A|B) = P(B) \]

   Or the probability of event A given that event B has occurred.
   \[ P(A|B) = P(A) \]

   Events may be either independent or dependent. For independent events, the occurrence of one event has no effect on the probability of occurrence of the second event.

   Examples of dependent events
   a) Your education
   b) Your income level
   a) Chicago Cubs win the National League
   b) Chicago Cubs win the World Series

   Examples of independent events
   a) Draw a jack of hearts from a full 52-card deck
   b) Draw a jack of clubs from a full 52-card deck
   a) Snow in Santiago, Chile
   b) Rain in Paris, France
Example for calculating the probability of statistically independent events:
A bucket contains 3 black balls and 7 green balls.
Draw a ball from the bucket, replace it, and draw a second ball.
1. The probability of a black ball drawn on the first draw is:
\[ P(B) = 0.3 \] (a marginal probability)
2. The probability of drawing a green ball on the first two draws:
\[ P(GG) = P(G) \times P(G) = 0.7 \times 0.7 = 0.49 \]

3. Counting Rules
This is not a standalone document. It is coordinated with your *Elementary Statistics* text. Triola (2007, pp. 179-186) discusses:

**Statistical Literacy and Critical Thinking**

- **Permutations and Combinations.** What is the basic difference between a situation requiring application of the permutations rule and one requiring the combinations rule? Combinations order doesn’t matter. Permutations order matters. (Permutations are ordered combinations, thus there are more possible permutations than combinations.)

- **Counting.** When trying to find the probability of winning the California Fantasy 5 lottery, it becomes necessary to find the number of different outcomes that can occur when 5 numbers between 1 and 39 are selected. Why can’t that number be found by simply listing all of the possibilities? (Too many)

- **Relative Frequency.** A researcher is analyzing a large sample of text in order to find the relative frequency of the word “top” (as in “top sales” or “top notch”) among three letter words. She wants to estimate the probability of getting the word “top” when a three letter word is randomly selected from a typical English text.

- **Probability.** Someone reasons that when a coin is tossed there are three possible outcomes: It comes up heads or tails, or it lands on its edge. With three outcomes on each toss, the fundamental counting rule suggests that there are nine possibilities (from \(3 \times 3 = 9\)) for two tosses of a coin.

Calculating Factorials, Combinations, and Permutations. Evaluate the given expressions and express all results using the usual format for writing numbers.

Factorials are noted by the “!” after the number. Thus, \(5!\) is pronounced five factorial.

\[ 5! = 5 \times 4 \times 3 \times 2 \times 1 = 120 \]

The number of permutations of \(n\) objects taken \(r\) at a time is given by the formula: \(P(n,r) = \frac{n!}{(n-r)!}\). Where \(n\) = number of objects taken \(r\) at a time.

\[ P(10,3) = \frac{10!}{(10-3)!} = \frac{10!}{7!} = 10 \times 9 \times 8 = 720 \text{ permutations} \]

The number of combinations of \(n\) objects taken \(r\) at a time is given by the formula: \(C(n,r) = \frac{n!}{r!(n-r)!}\)

\[ C(3,2) = \frac{3!}{2!(3-2)!} = \frac{6}{2} = 3 \]

This is not a standalone document. It is coordinated with your *Elementary Statistics* text. The following commentary is used to assist you in your decision making. Triola (2007, p. 182-185) states:

When we use the term permutations, arrangements, or sequences, we imply that order is taken into account in the sense that different orderings of the same items are counted separately. The letters \(ABC\) can be arranged six different ways: ABC, ACB, BAC, BCA, CAB, CBA. In the following example, we are asked to find the total number of different sequences that are possible. That suggests use of the permutations rule.
**Example:** Clinical Trial of New Drug - When testing a new drug, Phase I involves only 8 volunteers and the objective is to assess the drug's safety. To be very cautious, you plan to treat the 8 subjects in sequence so that any particularly adverse effect can allow for stopping the treatments before any other subjects are treated. If 10 volunteers are available, and 8 of them are to be selected, how many different sequences of 8 subjects are possible?

**Solution:** We have \( n = 10 \) different subjects available, and we plan to select \( r = 8 \) of them without replacement. The number of different sequences of arrangements is found as shown:

\[
nPr = \frac{n!}{(n-r)!} = \frac{10!}{(10-8)!} = 1,814,400
\]

There are 1,814,400 different possible arrangements (variety of sequences) of 8 subjects selected from the 10 that are available. The size of that result indicates that it is not practical to list the sequences or somehow consider each one of them individually.

**Requirements**
1. There are \( n \) items available and some items are identical to others.
2. We select all of the \( n \) items (without replacement).
3. We consider rearrangements of distinct items to be different sequences.

If the preceding requirements are satisfied and if there are \( n! \) alike, \( n_2! \) alike, ..., \( n_k! \) alike, the number of permutations (or sequences) of all items selected without replacement is:

\[
n! \div n! \div n_2! \div ... \div n_k!
\]

4. **Measures of Central Tendency and Distribution**

This is not a standalone document. It is coordinated with your *Elementary Statistics* text. Bluman (2007, p. 93-126) notes:

**Statistical Measures**
Purposeful analysis of mass data requires several kinds of statistical measures:

- Measures of central tendency
- Measures of relative position
- Measures of variability, spread, or dispersion
- Measures of relationship

There are other statistical measures, but they are beyond the scope of this review. The statistical symbols below may be different than other references you have viewed. Different statistical references use a variety of symbols for the terms. The table below shows the symbols for various test scores:

<table>
<thead>
<tr>
<th>Term</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>( \mu )</td>
</tr>
<tr>
<td>Median</td>
<td>Md</td>
</tr>
<tr>
<td>Mode</td>
<td>Mo</td>
</tr>
<tr>
<td>The sum of</td>
<td>( \Sigma )</td>
</tr>
<tr>
<td>Frequency</td>
<td>( f )</td>
</tr>
<tr>
<td>Number (of scores)</td>
<td>( N ) or ( n )</td>
</tr>
<tr>
<td>Raw scores</td>
<td>( X )</td>
</tr>
<tr>
<td>Percentile rank</td>
<td>PR</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>SD or ( \sigma )</td>
</tr>
</tbody>
</table>
Measures of Central Tendency

Measures of central tendency address the “average quantity” of a series of numbers, characteristics, or scores. A group of data is often described in terms of some measure near the middle of its distribution, which represents it. The three most commonly used measures are the mean, the median, and the mode.

Properties and Uses of the Concept of Central Tendency

The Mean (commonly called the average)
1. One computes the mean by using all the values of the data.
2. The mean varies less than the median or mode when samples are taken repeatedly from the same population, and all three measures are computed for these samples.
3. The mean is used in computing other statistics, such as the variance.
4. The mean for the data set is unique and not necessarily one of the data values.
5. The mean cannot be computed for an open-ended frequency distribution.
6. The mean is affected by extremely high or low values called outliers and may not be the appropriate average to use in all situations.

The Median (like the median in a road, it is near the middle)
1. The median is used when one must find the center or middle value of a data set.
2. The median is typically used when one must determine whether the data values fall into the upper half or lower half of the distribution.
3. The median is used for open-ended distributions.
4. The median is affected less than the mean by extremely high or extremely low values.

The Mode
1. The mode is used when the most typical case is desired.
2. The mode is the easiest average to compute (you simply count up the number of times a value occurs).
3. The mode can be used when the data are nominal such as religious preference, gender, or political affiliation.
4. The mode is not always unique unto itself. A data set can have more than one mode or the mode may not exist for a data set at all (all values may occur at the same rate).

The Midrange
1. The midrange is easy to compute.
2. The midrange gives the midpoint of a string of numbers.
3. The midrange can be affected by extreme values.

This is not a standalone document. It is coordinated with your *Elementary Statistics* text. Bluman (2007, p. 109) notes:

Uses of the Variance and Standard Deviation Concepts
1. Variances and standard deviations are used to determine the spread (what it looks like) of the data. If the variance or standard deviation is large, the data are said to be dispersed. This information is useful in comparing two or more data sets to determine which is more (most) variable (how consistently diverse their patterns are).

Mean variance and standard deviation are used to determine the consistency of a variable (whether the data are compacted together or not).

The variance and standard deviation are used to describe the number of data values that fall within specific intervals in a distribution. For example, Chebyshev's theorem shows that for any distribution that is normally distributed, at least 75% of the data values will fall within
two standard deviations (two unit intervals) around the mean.

Finally, the variance and standard deviation concepts are used quite often in inferential statistics work (drawing inferences from samples of data).

Whenever two samples use the same measures, the variance and standard deviation for each can be compared directly. For example, suppose an automobile dealer wanted to compare the standard deviation of miles driven for the motor bikes he received as trade-ins on new motor bikes scooters. He found that for a specific year, the standard deviation for Hondas was 422 miles and the standard deviation for Suzukis was 350 miles. He could say that the variation in mileage was greater in the Hondas.

What if a manager wanted to compare the standard deviations of two different variables like the number of sales per salesperson over a three-month period and the commissions made by these salespeople?

A statistic that allows one to compare standard deviations when the units are different, as in this example, is called the coefficient of variation.

The coefficient of variation is the standard deviation divided by the mean. The result is expressed as a percentage.

The mean of the number of sales of bikes over a three-month period is 87 and the standard deviation is 5. The mean of the commissions is $5,225 and the standard deviation is $773.

Compare the variations:
The coefficients of variation are:

\[
C_{\text{Var}} = \frac{87}{5} \times 100\% = 5.7\%
\]

\[
C_{\text{Var}} = \frac{5225}{773} \times 100\% = 14.8\%
\]

Since the coefficient of variation is larger for commissions, the commissions are typically more variable than the sales.

**Distributions**
This is not a standalone document. It is coordinated with your *Quantitative Analysis for Management* text. Render, Stair, & Hanna (2012, pp. 38-48) note:

**The Binomial Distribution**

Many business experiments can be characterized by the Bernoulli process. The Bernoulli process is described by the binomial probability distribution.

- Each trial has only two possible outcomes.
- The probability of each outcome stays the same from one trial to the next.
- The trials are statistically independent.
- The number of trials is a positive integer.

The binomial distribution is used to find the probability of a specific number of successes in \( n \) trials.
We need to know:

\[
\begin{align*}
\text{number of trials} & = n \\
\text{probability of success on any single trial} & = p
\end{align*}
\]
We let:

\[ r = \text{number of successes} \]
\[ q = 1 - p = \text{the probability of a failure} \]

The binomial formula is

\[
\text{Probability of } r \text{ successes in } n \text{ trials} = \frac{n!}{r!(n-r)!} p^r q^{n-r}
\]

The symbol \(!\) means factorial and \( n! = n(n-1)(n-2)(n-3)\ldots 1 \)

For example \( 4! = (4)(3)(2)(1) = 24 \)

By definition
\( 1! = 1 \) and \( 0! = 1 \)

Below is an example for the probability of the number of heads in 5 coin tosses. The probability is 50% on each toss, either heads or tails.

**Binomial Distribution for \( n = 5 \) and \( p = 0.50 \)**

<table>
<thead>
<tr>
<th>Number of Heads</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.03125</td>
</tr>
<tr>
<td>1</td>
<td>0.15625</td>
</tr>
<tr>
<td>2</td>
<td>0.31250</td>
</tr>
<tr>
<td>3</td>
<td>0.31250</td>
</tr>
<tr>
<td>4</td>
<td>0.15625</td>
</tr>
<tr>
<td>5</td>
<td>0.03125</td>
</tr>
</tbody>
</table>

Typically, we do not use the formula to solve binomial problems because using the binomial table is easier. The binomial table is available on the web from many sources. Download one for your convenience. To use the table, look up the \( n \) (the number of trials). The vertical axis is \( r \) (the number of successes). The horizontal axis is \( P \) (the probability of a success).

For example, a company makes a circuit board.
- Every hour a random sample of 5 circuit boards is taken.
- The probability of one circuit board being defective is 0.15.
- What is the probability of finding 3 or 4 defective?
- Below is a portion of a binomial table:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( r )</th>
<th>P ( 0.05 )</th>
<th>P ( 0.10 )</th>
<th>P ( 0.15 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0.7738</td>
<td>0.5905</td>
<td>0.4437</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0.2036</td>
<td>0.3281</td>
<td>0.3915</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.0214</td>
<td>0.0729</td>
<td>0.1382</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.0011</td>
<td>0.0081</td>
<td>0.0244</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>0.0000</td>
<td>0.0005</td>
<td>0.0022</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Call a defect a “success” because you are looking for failures.
From the table, the probability of 3 defects is 0.0244 and the probability of 4 defects is 0.0022.
5. The normal distribution

- The normal distribution is symmetrical with the high midpoint representing the mean.
- Shifting the mean does not change the shape of the distribution.
- Values on the X axis are measured in the number of standard deviations away from the mean.
- As the standard deviation becomes larger, the curve flattens.
- As the standard deviation becomes smaller, the curve becomes steeper.
- The curve is composed of three different curves. The two on each end look like smiles, and the one in the center looks like a frown. You can remember that by smiling twice as much as you frown.
- On each side, where the smile meets the frown is a special point. That point is exactly one standard deviation away from the mean.
- About 68% (closer to 68.268%) of the data falls between 1 standard deviation below the mean and 1 standard deviation above the mean. (See the graph below.)
- About 95% (closer to 95.45%) of the data falls between 2 standard deviations below the mean and 2 standard deviations above the mean.
- About 99.7% of the data falls between 3 standard deviations below the mean and 3 standard deviations above the mean.

Just like the binomial distribution, we typically use the standard normal table to find the probability of a value being less than a certain number. Using a standard normal table is a two-step process.

**Step 1** Convert the normal distribution into a standard normal distribution.
- A standard normal distribution has a mean of 0 and a standard deviation of 1.
- The new standard random variable is \( z \).

\[
Z = \frac{X - \mu}{\sigma}
\]

Where:
- \( X \) = value of the random variable we want to measure
- \( \mu \) = mean of the distribution
- \( \sigma \) = standard deviation of the distribution
- \( Z \) = number of standard deviations from the \( X \) to the mean, \( \mu \)

For example, \( \mu = 100 \), \( \sigma = 15 \) and we want to find the probability that \( X \) is less than 130

\[
Z = \frac{X - \mu}{\sigma} = \frac{130 - 100}{15} = \frac{30}{15} = 2 \text{ std dev}
\]
Step 2 Look up the probability from a table of normal curve areas (from the web). The column on the left has the Z values. The row at the top has the second decimal point for the Z values.

### AREA UNDER THE NORMAL CURVE

<table>
<thead>
<tr>
<th>Z</th>
<th>0.00</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.8</td>
<td>0.96407</td>
<td>0.96485</td>
<td>0.96562</td>
<td>0.96638</td>
</tr>
<tr>
<td>1.9</td>
<td>0.97128</td>
<td>0.97193</td>
<td>0.97257</td>
<td>0.97320</td>
</tr>
<tr>
<td>2.0</td>
<td>0.97725</td>
<td>0.97784</td>
<td>0.97831</td>
<td>0.97882</td>
</tr>
<tr>
<td>2.1</td>
<td>0.98214</td>
<td>0.98257</td>
<td>0.98300</td>
<td>0.98341</td>
</tr>
<tr>
<td>2.2</td>
<td>0.98610</td>
<td>0.98645</td>
<td>0.98679</td>
<td>0.98713</td>
</tr>
</tbody>
</table>

\[ P(X < 130) = P(Z < 2.00) = 0.97725 \]

Thus, the probability of \( X < 130 \) is 0.97725.

6. Sampling and Estimation
   The following is from [https://sites.google.com/site/fundamentalstatistics/chapter-9](https://sites.google.com/site/fundamentalstatistics/chapter-9), accessed May 15, 2014.
   
   **Using Samples to Estimating the of Population**
In a typical research scenario, a researcher randomly selects a sample of individuals from a larger population and then uses that sample to make inferences about the population. A population is a collection of all individuals or all scores that a researcher wants to make an inference about. Populations are theoretically infinitely large; hence, it is impossible to know most population parameters. Indeed, this is the reason for collecting and using samples and making inferences about populations based on sample data. Populations are just too inconvenient (humans like things that are convenient, like a coffee maker in your office). Thus, researchers use sample data to estimate the population parameters and infer what would be found in the population.

How good of estimators are sample statistics; that is, how accurately does a sample mean estimate a population mean (the same question also pertaining to other parameters)? In this chapter, we go through how the sample mean is an unbiased estimator of the population mean ($\mu$), but sample variance ($s^2$) and standard deviation ($s$) are biased estimators of the true population variance ($\sigma^2$) and standard deviation ($\sigma$), respectively, and how that bias is corrected.

The Sample Mean as an Unbiased Estimator of the Population Mean
Assume you are a statistics professor and one of your classes is populated with $N = 25$ students. On a quiz, which has a range of 0 to 10 points, you obtain the scores in the table below. Below the scores in the table are this population's (class) mean ($\mu$), variance ($\sigma^2$), and standard deviation ($\sigma$). For this discussion, assume you do not know these three population parameters.

<table>
<thead>
<tr>
<th>Statistics Quiz Scores ($N = 25$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

$\mu = 5.400$, $\sigma^2 = 7.040$, $\sigma = 2.636$

Say you meet with other statistics teachers to compare grades. You do not know the population values of your quiz data (maybe you do not have time to calculate the parameters for all 25 students), so you randomly sample $n = 5$ quiz scores and calculate the sample mean, sample variance ($s^2$), and sample standard deviation ($s$). This sample contains the following scores: 10, 8, 5, 5, and 3; hence, the sample mean is 6.2, which is not equal to the population mean of 5.400. Although you do not know that this sample's mean differs from the population by +0.8 (6.2 – 5.4), this difference is not a big problem. The difference between this sample's mean and the population mean is called sampling error, which results from a sample not perfectly reflecting the population. That is, any time that you select a sample from a population, you are not selecting all of the variability and individual differences in that population; hence, the sample mean might very well be different from the population mean. Indeed, the only way that a sample can perfectly represent a population is if the sample included the entire population, which in that case you would have the population!

Say that you randomly sample another 5 scores (9, 6, 5, 4, and 2) and find this second sample's mean is 5.2. In this case, the sampling error is -0.2; thus, this sample's mean is less than the population mean. If you were to continue with this little exercise and generate more samples from this population, you would find that some sample means are greater than $\mu$, some sample means are less than $\mu$, and many sample means are equal to $\mu$. More importantly, this indicates that sample means will never be consistently greater than or consistently less than the population mean; thus, the sample mean is an unbiased estimator of $\mu$.

7. Hypothesis Testing
The following is from https://sites.google.com/site/fundamentalstatistics/chapter-14, accessed May 15, 2014
All inferential statistics boil down to a **test-statistic**, which is a ratio of the size of a relationship, or the **effect**, to some measurement of random sampling **error**. Thus, any test statistic can be summarized as:

\[
test = \frac{effect}{error}
\]

We generally want the ratio of effect to error to be large. The question addressed in this and the following chapters is how large that ratio needs to be for a relationship to be **statistically significant**. But first, remember that in any study one must explicitly state the predictions of his or her null and alternate hypotheses. Recall that the **null hypothesis** (H₀) is the prediction that the expected relationship will not be observed and the **alternative hypothesis** (H₁) is the prediction that the expected relationship will be observed in the sample data. In the ginko-baloba example above, one might predict that taking ginko-baloba, which some believe affects memory, will affect student GPAs. In this case, the null hypothesis would predict that taking ginko-baloba will not lead to a difference in GPA compared to taking a placebo (expected outcome not observed); and the alternate hypothesis would predict that taking ginko-baloba will lead to a difference in GPA compared to taking a placebo (expected outcome observed). When stating hypotheses, symbols are used and the null and alternate hypotheses should be stated in terms of their populations. For example, in the ginko-baloba example above, the null and alternate hypotheses would be written symbolically as (be sure to use subscripts to distinguish the two populations):

\[
H₀: \mu_{Ginko} = \mu_{Placebo}
\]

\[
H₁: \mu_{Ginko} \neq \mu_{Placebo}
\]

Both hypotheses are symbolically stating the predictions described in the preceding paragraph. That is, the null hypotheses states, symbolically, that taking ginko-baloba will not lead to a change in GPA; or there will be no difference in mean GPA between the students taking ginko-baloba and the students taking the placebo. In contrast, the alternative hypothesis is stating symbolically that there will be a difference between mean GPA between the ginko-baloba group and the placebo group. The point is that when stating hypotheses, you should write them symbolically rather than verbally, because meanings of symbolic statements are universal, whereas words could be misinterpreted.

You may wonder why population symbols are used rather than sample symbols. Remember, inferential statistics use sample data to make inferences about what should be found in the population from which the sample came; thus, hypotheses should reflect the inferences, which is why population symbols are used.

**Null Hypothesis Significance Testing (NHST)**

The logic to null hypothesis significance testing (NHST) may seem a little strange, convoluted, and/or flawed at first. Indeed, there are some scientists who are critical of hypothesis testing and believe it should be banned, a point that I tend to agree with. In hypothesis testing, the null hypothesis, not the alternate hypothesis, is tested. Of course, this seems strange, because the alternate hypothesis is usually the prediction one wants to confirm; hence, it would seem logical to test the alternate hypothesis. But if this was so, then a researcher would be seeking evidence only to confirm predictions and, thus, may be biased in evidence gathering. In scientific
reasoning, you want to disconfirm objections to your prediction; thus, one should seek to disconfirm his or her null hypothesis. If you can disconfirm the null, you can use the alternate. (I know, this already sound's weird!)

In null hypothesis testing, you start out assuming the null hypothesis is true. You then collect data and use inferential statistics to calculate the probability of observing that data given that the null hypothesis is true. That is, inferential statistics are used to tell us $p(\text{Data} \mid \text{H}_0 = \text{True})$. If the probability of observing the data given that the null is true is very low, which means that observing the data is unlikely assuming that the null hypothesis is true, then one "rejects" the null hypothesis, because it is unlikely to be true. Stated differently, 'if the probability of observing the data is very low (assuming the null is true), then the null is unlikely to be correct, and we reject the null as being correct'. Once the null hypothesis is rejected, one can accept the alternate hypothesis as the next best option.

But at what point is the probability of observing data low enough that they are unlikely to be observed if the null hypothesis was indeed true? Again, you want the probability of observing the data to be low, not high. The first question that may pop into your head is, 'why low?' Remember, as stated in the preceding paragraph, you start out assuming that the null hypothesis is true. If the probability of observing your data is very low, this suggests that this assumption is unlikely to be true. Thus, if there is a small probability of observing the data if the null is true, it suggests the null is unlikely to be true; therefore, it can be rejected.

So how low does the probability of an outcome have to be before we can say that the null hypothesis is unlikely to be true and it is rejected? The accepted level of statistical significance is $p = .05$ or less, which is called your alpha-level ($\alpha$). That is, assuming the null hypothesis is true, if the probability of an event (the outcome of an inferential test) is $p = .05$ or less, then it is unlikely the null hypothesis is true and it can be rejected. In such cases, we conclude that the outcome is statistically significant. The 'p-Value' one chooses to use is the $\alpha$ level. Researchers use different p-Values for different reasons, but the generally accepted level of statistical significance is $\alpha = .05$ or less. Important: The $p$-Value is not the probability that the null hypothesis is true; it is the probability of observing data given the null hypothesis is true.

From the ginko-baloba example in the preceding section, say we find the mean GPA for students taking ginko-balboa is 3.5 and the mean GPA for students taking the placebo is 3.2. Using the correct test, say the probability of this difference is less than $\alpha = .05$. In this case, we would conclude there is a statistically significant difference in GPA between students taking ginko-baloba and students taking a placebo; we reject the null hypothesis and accept the alternate hypothesis and conclude that taking ginko-baloba as a freshman college student increases overall GPA.

**Type I vs. Type II Errors**

Is hypothesis testing perfect? Is the outcome of a study going to be found each and every time a study is replicated? Sadly, the answer is no. It is possible to make a mistake in hypothesis testing, because hypothesis testing is based on probabilities which are not 'absolute certainties.' Remember, whatever the $\alpha$ level is that you select, this is the probability of observing the outcome that you did if the null hypothesis is true. Thus, if $\alpha = .05$, it means that there is a very small chance that you will observe the outcome with the null being true; but it also means that there is a .05 chance that the null is true given this outcome. That is, there is a small chance that you will observe this data, even with the null being true. In short, because hypothesis testing is based on probabilities of being correct or incorrect (rejecting vs. retaining the null), there is always going to be some probability that your decision can be wrong. Indeed, there are two incorrect decisions (statistics errors) that can be made in hypothesis testing.
Before describing these statistical errors, first envision the possible decisions that can be made about the outcome of hypothesis testing and the possible “true nature of the world.” In hypothesis testing, if the obtained value (test statistic) is greater than or equal to the critical value, the null hypothesis is rejected; but if the obtained value is less than the critical value, the null hypothesis is retained. Thus, there are two possible decisions that can be made about a hypothesis, and making one depends on the test statistic being statistically significant. In reality (i.e., in a population) the test statistic may be significant or not. Remember, in hypothesis testing you are using sample data to make inferences about what you would expect to find in a population. It is possible that you conclude a sample mean is statistically different from a population mean, but in reality there is no difference. There are four possible combinations of your statistical decision regarding the outcome of a hypothesis test and the true nature of the world, which can be seen in the contingency table below:

<table>
<thead>
<tr>
<th>True Nature of the World</th>
<th>Decision Based on Outcome of Hypothesis Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₀ Actually True</td>
<td>Reject H₀</td>
</tr>
<tr>
<td></td>
<td>Type I Error ( p = \alpha )</td>
</tr>
<tr>
<td>H₀ Actually False</td>
<td>Correct Decision ( p = 1 - \beta )</td>
</tr>
<tr>
<td></td>
<td>Fail to Reject H₀</td>
</tr>
<tr>
<td></td>
<td>Correct Decision ( p = 1 - \alpha )</td>
</tr>
</tbody>
</table>

From the table, you can see that there are two possible combinations that lead to a correct decision and two possible combinations that lead to decision errors. I’ll focus on the errors:

A **Type I Error** is committed when the null hypothesis is rejected, but in the true nature of the world, the null hypothesis is actually true and should be retained. The probability of committing this error is equal to the selected \( \alpha \) level. To decrease the chance of making a Type I Error, a smaller \( \alpha \) can be selected (e.g., .01, .001, .0001, etc.) and used as a basis for retaining/rejecting the null. Using the ginko-baloba example, the freshmen population’s mean GPA was \( \mu = 2.8 \) and our sample mean was \( \alpha = 3.2 \). Because the z-Test was significant, the null hypothesis was rejected. But what if a few other freshmen in the population, who had not been included in the sample, were also taking ginko-baloba? Say we went back and surveyed all of the freshmen students at the university to identify students taking ginko-baloba, including our sample. Say the mean GPA of the population of freshmen taking ginko-baloba is \( \mu = 2.9 \). All things being equal, this population mean and the population mean of all freshmen (\( \mu = 2.8 \)) would not significantly differ. Thus, in reality, there is no statistical difference in GPA between students taking ginko-baloba and those not taking ginko-baloba. Thus, our initial decision to reject the null hypothesis was a Type I error.

The second type of error, a **Type II Error**, is committed when you fail to reject the null hypothesis (you retain the null); but in the true nature of the world, the null hypothesis is false and should be rejected. The probability of making a Type I Error is equal to something called \( \beta \) (\( \beta \)). The probability of not making a Type II Error (correctly rejecting the null hypothesis) is equal to \( 1 - \beta \), which is the **power** a statistical test has to detect a significant difference. This can be illustrated using the non-significant difference between the two populations’ means alluded to in the preceding paragraph. Say we gathered data across several universities looking for freshmen taking ginko-baloba vs. students not taking ginko-baloba and found a significant difference. Thus, in the true nature of the world, there is a significant difference in GPA between students taking ginko-baloba and students not taking ginko-baloba. Hence, the decision to retain the null hypothesis based on data from only one university would have been a Type II error.

8. **Correlation and Regression**

This is not a standalone document. It is coordinated with your *Quantitative Analysis for Management* text. Render, Stair, & Hanna (2012, pp. 116-121) note:
Correlation is the relationship between variables. See the figure that shows what the correlation coefficient means. Regression is an analysis tool to determine the correlation between two variables.

Regression analysis is a valuable tool for managers. Regression can be used to understand the relationship between variables and predict the value of one variable based on another variable. Simple linear regression has only two variables. Multiple regression models have more variables. The variable to be predicted is called the dependent variable, sometimes called the response variable. The value of this variable depends on the value of the independent variable, sometimes called the predictor variable or explanatory variable. A scatter diagram or scatter plot should be your first step in investigating the relationship between two variables. The independent variable is normally plotted on the X-axis and the dependent (response) variable on the Y-axis.

Regression models are used to test if there is a relationship between variables. There is some random error, \( E \), that can’t be predicted. The formula for the line regression equation is:

\[
Y = B_0 + B_1X + \epsilon
\]

Where:
- \( Y \) = dependent (response) variable
- \( X \) = independent (predictor or response) variable
- \( B_0 \) = intercept (value of Y when \( X = 0 \))
- \( B_1 \) = the slope of the intercept line
- \( \epsilon \) = random error

True values for the slope and the intercept are not known, so they are estimated from the sample data.

\[
\hat{Y} = b_0 + b_1X
\]

Where:
- \( \hat{Y} \) = predicted value of \( Y \)
- \( b_0 \) = estimate of \( B_0 \), based on sample results
- \( b_1 \) = estimate of \( B_1 \), based on sample results

What the correlation coefficient means:
Example
The sales manager of a large apartment rental complex feels the demand for apartments may be related to the number of newspaper ads placed during the previous month. She has collected the data shown in the accompanying table.

<table>
<thead>
<tr>
<th>Ads purchased, (X)</th>
<th>Apartments leased, (Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>40</td>
<td>16</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>25</td>
<td>9</td>
</tr>
<tr>
<td>15</td>
<td>10</td>
</tr>
<tr>
<td>35</td>
<td>16</td>
</tr>
</tbody>
</table>

We can find a mathematical equation by using the least squares regression approach.
(Note: Round-off error may cause this to be slightly different from a calculator solution.)

\[
\begin{align*}
\sum Y &= 80 \\
\sum X &= 184 \\
\sum (X - \bar{X})^2 &= 774 \\
\sum (X - \bar{X})(Y - \bar{Y}) &= 306 \\
\end{align*}
\]

\[
\begin{align*}
\bar{Y} &= \frac{80}{8} = 10; \\
\bar{X} &= \frac{184}{8} = 23 \\

b_1 &= \frac{306}{774} = 0.395 \\

b_0 &= 10 - 0.395(23) = 0.915 \\

Thus Y = 0.395X + 0.915
\]

9. Time-series forecasting
This is not a standalone document. It is coordinated with your *An Introduction to Management Science* text. Anderson, Sweeney, & Williams (2008, pp. 713-716) state:

Time series: A set of observations measured at successive points in time or over successive periods of time.
Forecast: A projection or prediction of future values of a time series.

Trend: The long-run shift or movement in the time series observable over several periods of data.

Cyclical component: The component of the time series model that results in periodic above-trend and below-trend behavior of the time series lasting more than one year.

Seasonal component: The component of the time series model that shows a periodic pattern over 1 year or less.

Irregular component: The component of the time series model that reflects the random variation of the actual time series values beyond what can be explained by the trend, cyclical, and seasonal components.

Moving averages: A method of forecasting or smoothing a time series by averaging each successive group of data points.

Mean squared error (MSE): An approach to measuring the accuracy of a forecasting model. This measure is the average of the squared differences between the actual time series values and the forecasted values.

Weighted moving averages: A method of forecasting or smoothing a time series by computing a weighted average of past time series values. The sum of the weights must equal 1.

Exponential smoothing: A forecasting technique that uses a weighted average of past time series values to arrive at smoothed time series values that can be used as forecasts.

Smoothing constant: A parameter of the exponential smoothing model that provides the weight given to the most recent time series value in the calculation of the forecast.

Mean absolute deviation (MAD): A measure of forecast accuracy. The average of the absolute values of the forecast errors.

Seasonal index: A measure of the seasonal effect on a time series. A seasonal index greater than 1 indicates a positive effect, a seasonal index of 1 indicates no seasonal effect, and a seasonal index less than 1 indicates a negative effect.

Deseasonalized time series: A time series that has had the effect of season removed by dividing each original time series observation by the corresponding seasonal index.

Regression analysis: A statistical technique that can be used to develop a mathematical equation showing how variables are related.

Causal forecasting methods: Forecasting methods that relate a time series to other variables that are believed to explain or cause its behavior.

Autoregressive model: A time series model that uses a regression relationship based or historical time series values to predict the future time series values.

Smoothing methods are easy to use and generally provide a high level of accuracy for short-range forecasts such as a forecast for the next time period. One of the methods, exponential smoothing, has minimal data requirements and is a good method to use when forecasts are required for large numbers of items.

Moving Averages
The moving averages method uses the average of the most recent n data values in the time series as the forecast for the next period.

\[ \text{Moving average} = \frac{\text{(the most recent } n \text{ data values})}{n} \]
The term *moving* indicates that as a new observation becomes available for the time series, it replaces the oldest observation in equation and a new average is computed. As a result, the average will change, or move, as new observations become available.

To illustrate the moving averages method, consider the 12 weeks of data presented. These data show the number of gallons of gasoline sold by a gasoline distributor over the past 12 weeks. The time series appears to be stable over time. In this case, smoothing methods are applicable.

### Gasoline Sales Times Series

<table>
<thead>
<tr>
<th>Week</th>
<th>Sales (1,000s of gallons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
</tr>
</tbody>
</table>

To use moving averages to forecast gasoline sales, we must first select the number of data values to be included in the moving average. For example, let us compute forecasts using a 3 week moving average. The moving average calculation for the first 3 weeks of the gasoline sales time series is:

\[
\text{Moving average (weeks 1-3)} = \frac{17+21+19}{3} = 19
\]

We then use this moving average value as the forecast for week 4. The actual value observed in week 4 is 23, so the forecast error in week 4 is 23 - 19 = 4. In general, the error associated with a forecast is the difference between the observed value of the time series and the forecast.

The calculation for the second 3 week moving average is:

\[
\text{Moving Average (weeks 2-4)} = \frac{21+19+23}{3} = 21
\]
Anderson, Sweeney, & Williams (2008) state:

Calculating the Seasonal Indexes

The figure below indicates that sales are lowest in the second quarter of each year, followed by higher sales levels in quarters 3 and 4. Thus, we conclude that a seasonal pattern exists for television set sales.

Graph of Quarterly Television Set Sales Time Series
Table Quarterly Data for Television Set Sales

<table>
<thead>
<tr>
<th>Year</th>
<th>Quarter</th>
<th>Sales (1,000s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4.8</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5.8</td>
</tr>
<tr>
<td>2</td>
<td>5.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7.4</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6.0</td>
</tr>
<tr>
<td>2</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>6.3</td>
</tr>
<tr>
<td>2</td>
<td>5.9</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8.4</td>
<td></td>
</tr>
</tbody>
</table>


This is not a standalone document. It is coordinated with your text *An Introduction to Management Science*. Anderson, Sweeney, & Williams (2008) state:

**Other types of forecasting methods**: In the preceding sections, we discussed several types of quantitative forecasting methods. Most of these techniques require historical data on the variable of interest, so they cannot be applied when no historical data are available. Furthermore, even when such data are available, a significant change in environmental conditions affecting the time series may make the use of past data questionable in predicting future values of the time series. Qualitative forecasting techniques offer an alternative in these and other cases. Some of these techniques are listed below:

**Delphi Method**

The Delphi method, originally developed by a research group at the Rand Corporation, attempts to develop forecasts through "group consensus." Members of a panel of experts, all of whom are both physically separated from and unknown to each other, are asked to respond to a series of questionnaires. Responses are tabulated and used to prepare a second questionnaire that contains information and opinions of the entire group. Respondents are asked to reconsider and possibly revise their previous responses. This process continues until the coordinator feels that some degree of consensus has been reached. The goal is to produce a narrow spread of opinions within which the majority of experts concur.

**Expert Judgment**

Qualitative forecasts often are based on the judgment of a single expert or represent the consensus of a group of experts. The experts individually consider information they believe will influence a group (e.g., the stock market and interest rates); then the experts combine their conclusions into a forecast. No formal model is used. Expert judgment is a forecasting method that is recommended when conditions in the past are not likely to hold in the future. Despite the fact that no formal quantitative model is used, expert judgment provides good forecasts in many situations.

**Scenario Writing**
Scenario writing consists of developing a conceptual scenario of the future that is based on a well-defined set of assumptions. Different assumptions lead to different scenarios. Management is to decide how likely each scenario is and decide accordingly.

10. Statistical Concepts in Quality Control
This is not a standalone document. It is coordinated with your Quantitative Analysis for Management text. Render, Stair, & Hanna (2012, pp. 603-613) note:

**Quality** is usually a major issue in a purchasing comparison, because poor quality can be expensive for both the firm and the customer. Quality of a product is the degree which the product or service meets specifications. **Quality control (QC)** is critical for the success of any organization.

In highly repetitive production situations, **statistical process control (SPC)** is important. **SPC** is the process of establishing and monitoring standard measurements and taking corrective action as a product or service is produced. Process output is measured by samples, and if the sample results fall outside specific ranges, the process is stopped and corrected.

A control chart is a graphical representation of data over time and shows the upper and lower limits of the parameter we want to control. Control charts are built using averages of small samples. The purpose of a control chart is to distinguish between natural and variations due to assignable causes.

Two of the most common control charts are the Mean and the Range Charts. The Mean is also called the $x$ - Chart and the Range is $r$ – Chart.

To construct these charts, you take multiple samples of the output from the process. The sample size is typically between 2 and 25 units of production, such as boxes filled with cereal.

- For each sample, you calculate the average and the range.
- Next, calculate the average of the average measurement from the samples and the average of the ranges.
- The charts are then based on factors: $A_2$ for the Average value, $D_4$ for the Upper Range, and $D_3$ for the Lower Range.

Example 16-11.
Small boxes of NutraFlakes cereal are labeled “net weight 10 ounces.” Each hour, random samples of size $n = 4$ boxes are weighed to check process control. Five hours of observations yielded the following:

<table>
<thead>
<tr>
<th>Time</th>
<th>Box 1</th>
<th>Box 2</th>
<th>Box 3</th>
<th>Box 4</th>
<th>Average</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>9 A.M.</td>
<td>9.8</td>
<td>10.4</td>
<td>9.9</td>
<td>10.3</td>
<td>10.10</td>
<td>0.60</td>
</tr>
<tr>
<td>10 A.M.</td>
<td>10.1</td>
<td>10.2</td>
<td>9.9</td>
<td>9.8</td>
<td>10.00</td>
<td>0.40</td>
</tr>
<tr>
<td>11 A.M.</td>
<td>9.9</td>
<td>10.5</td>
<td>10.3</td>
<td>10.1</td>
<td>10.20</td>
<td>0.60</td>
</tr>
<tr>
<td>12 P.M.</td>
<td>9.7</td>
<td>9.8</td>
<td>10.3</td>
<td>10.2</td>
<td>10.00</td>
<td>0.60</td>
</tr>
<tr>
<td>1 P.M.</td>
<td>9.7</td>
<td>10.1</td>
<td>9.9</td>
<td>9.9</td>
<td>9.90</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Calculate the average of the average; the answer is 10.4 ounces
Calculate the average of the range; the answer is 0.52 ounces
Next, look in Table 16-2 for the values of the factors: $A_2$ for the Average value, $D_4$ for the Upper Range, and $D_3$ for the Lower Range.
\( n = 4 \) from Table 16.2; \( A_2 = 0.729; D_4 = 2.282; \ D_3 = 0. \)

\[
\begin{align*}
UCL_{\bar{X}} &= \bar{X} + A_2 \times \bar{R} = 10.04 + 0.729 \times 0.52 = 10.42 \\
LCL_{\bar{X}} &= \bar{X} - A_2 \times \bar{R} = 10.04 - 0.729 \times 0.52 = 9.66 \\
UCL_R &= D_4 \times \bar{R} = 2.282 \times 0.52 = 1.187 \\
LCL_R &= D_3 \times \bar{R} = 0 \times 0.52 = 0
\end{align*}
\]

The smallest sample mean is 9.9 and the largest 10.2. Both are well within the control limits. The largest range is 0.6, which is well under the upper control limit. Thus, we can conclude that the process is currently within control.

Table 16.2 Factors for computing control chart limits

<table>
<thead>
<tr>
<th>Sample Size, ( n )</th>
<th>Mean Factor, ( A_2 )</th>
<th>Upper Range, ( D_4 )</th>
<th>Lower Range, ( D_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.880</td>
<td>3.268</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1.023</td>
<td>2.574</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0.729</td>
<td>2.282</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0.557</td>
<td>2.115</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0.483</td>
<td>2.004</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0.419</td>
<td>1.924</td>
<td>0.076</td>
</tr>
<tr>
<td>8</td>
<td>0.373</td>
<td>1.864</td>
<td>0.136</td>
</tr>
<tr>
<td>9</td>
<td>0.337</td>
<td>1.816</td>
<td>0.184</td>
</tr>
<tr>
<td>10</td>
<td>0.308</td>
<td>1.777</td>
<td>0.223</td>
</tr>
<tr>
<td>12</td>
<td>0.266</td>
<td>1.716</td>
<td>0.284</td>
</tr>
<tr>
<td>14</td>
<td>0.235</td>
<td>1.671</td>
<td>0.329</td>
</tr>
<tr>
<td>16</td>
<td>0.212</td>
<td>1.636</td>
<td>0.364</td>
</tr>
<tr>
<td>18</td>
<td>0.194</td>
<td>1.608</td>
<td>0.392</td>
</tr>
<tr>
<td>20</td>
<td>0.180</td>
<td>1.586</td>
<td>0.414</td>
</tr>
<tr>
<td>25</td>
<td>0.153</td>
<td>1.541</td>
<td>0.459</td>
</tr>
</tbody>
</table>


**Part B. Quantitative Operations Management Techniques**

1. **Linear programming**

This is not a standalone document. It is coordinated with your text *An Introduction to Management Science*. Anderson, Sweeney, & Williams (2008, pp. 32-46) suggest:

**A Maximization Problem**

Par, Inc. is a small manufacturer of golf equipment and golf supplies whose management has decided to move into the market for medium and high-priced golf bags. After a thorough investigation of the steps involved in manufacturing a golf bag, management determined that each golf bag produced will require:

1. Cutting and dyeing the material
2. Sewing
3. Finishing
4. Inspection and packaging

The director of manufacturing analyzed each of the operations and concluded that if the company produces a medium-priced standard model, each bag will require 7/10 hour in the cutting and
dyeing department, 1/2 hour in the sewing department, 1 hour in the finishing department, and 1/10 hour in the inspection and packaging department. The more expensive deluxe model will require 1 hour for cutting and dyeing, 5/6 hour for sewing, 2/3 hour for finishing, and 1/4 hour for inspection and packaging.

Par's production is constrained by a limited number of hours available in each department. After studying departmental workload projections, the director of manufacturing estimates 630 hours for cutting and dyeing, 600 hours for sewing, 708 hours for finishing, and 135 hours for inspection and packaging will be available for the production of golf bags during the next three months.

The accounting department assigned all relevant variable costs and arrived at prices for both bags that will result in a profit contribution of $10 for every standard bag and $9 for every deluxe bag produced.

Anderson, Sweeney, & Williams (2008, pp. 35-37) note:

Mathematical Model for the Par problem:

\[
\text{Max } 10S + 9D \\
\text{subject to (s.t.)} \\
\frac{7}{10}S + 1D \leq 630 \quad \text{Cutting and dyeing} \\
\frac{1}{2}S + \frac{5}{6}D \leq 600 \quad \text{Sewing} \\
S + \frac{2}{3}D \leq 708 \quad \text{Finishing} \\
\frac{1}{10}S + \frac{1}{4}D \leq 135 \quad \text{Inspection and packaging} \\
S, D \geq 0
\]

Find the product mix (i.e., the combination of S and D) that satisfies all the constraints and, at the same time, yields a value for the objective function that is greater than or equal to the value given by any other feasible solution. Once these values are calculated, we will have found the optimal solution to the problem.

This mathematical model of the Par problem is a linear programming model, or linear program. The problem has the objective and constraints that are common properties of all linear programs. This is linear, as your decision rule is the best alternative of more than one alternative (doable/feasible) choice.

Mathematical functions in which each variable appears in a separate term and is raised to the first power are called linear functions. The objective function (10S + 9D) is linear because each decision variable appears in a separate term and has an exponent of 1. The amount of production time required in the cutting and dyeing department (7/10S + 1D) is also a linear function of the decision variables for the same reason. Similarly, the functions on the left-hand side of all the constraint inequalities (the constraint functions) are linear functions. Thus, the mathematical formulation of this problem is referred to as a linear program.

Remember, linear programming has nothing to do with computer linear programming.

The use of the word programming here means "choosing a course of action." Linear programming involves choosing a course of action when the mathematical model of the problem contains only linear functions. There is, then, a best approach or choice.
The following definitions are both answers and methods of actions of calculating alternatives subject to the final goal.

Constraint: An equation or inequality that rules out certain combinations of decision variables as feasible solutions.

Problem formulation: The process of translating the verbal statement of a problem into a mathematical statement called the mathematical model.

Decision variable: A controllable input for a linear programming model.

Non-negativity constraints: A set of constraints that requires all variables to be non-negative.

Mathematical model: A representation of a problem where the objective and all constraint conditions are described by mathematical expressions.

Linear programming model: A mathematical model to include a linear objective function, a set of linear constraints, and non-negative variables. Choice of best answer when there are more than one correct answers/options.

Linear program: Another term for linear programming model.

Linear functions: Mathematical expressions in which the variables appear in separate terms and are raised to the first power.

Feasible solution: A solution that satisfies all the constraints.

Feasible region: The set of all feasible solutions.

Slack variable: A variable added to the left-hand side of a less-than-or-equal-to constraint to convert the constraint into an equality. The value of this variable can usually be interpreted as the amount of unused resources.

Standard form: A linear program in which all the constraints are written as equalities. The optimal solution of the standard form of a linear program is the same as the optimal solution of the original formulation of the linear program.

Redundant constraint: A constraint that does not affect the feasible region. If a constraint is redundant, it can be removed from the problem without affecting the feasible region.

Extreme point: Graphically speaking, extreme points are the feasible solution points occurring at the vertices or "corners" of the feasible region. With two-variable problems, extreme points are determined by the intersection of the constraint lines.

Surplus variable: A variable subtracted from the left-hand side of a greater-than-or-equal-to constraint to convert the constraint into an equality. The value of this variable can usually be interpreted as the amount over and above some required minimum level.

Alternative optimal solutions: The case in which more than one solution provides the optimal value for the objective function.
Infeasibility: The situation in which no solution to the linear programming problem satisfies all the constraints.

Unbounded: If the value of the solution may be made infinitely large in a maximization linear programming problem or infinitely small in a minimization problem without violating any of the constraints, the problem is said to be unbounded.

2. Project Scheduling (including PERT and CPM)

This is not a standalone document. It is coordinated with your text An Introduction to Management Science. Anderson, et al. (2008, pp. 457-494) state:

PERT (Program Evaluation Review Technique) is used to plan the scheduling of A, individual activities that make up a project.

A PERT network can be constructed to model the precedence (comes before) of the activities. The nodes of the network represent the activities towards conclusion of the project.

PERT can be used to determine the earliest/latest start and finish times for each activity, the entire project completion time, and the slack time for each activity.

“it” defines the initial number through and including the last number.

“it” DEFINES THE INITIAL NUMBER THROUGH AND INCLUDING THE LAST NUMBER

A critical path for the network is a path consisting of activities with zero slack (no slack time).

In the three time estimate approach, the time to complete an activity is assumed to follow a Beta distribution. Its mean is $t = (a + 4m + b)/6$ and its variance = $(b-a)/6)^2$. Here $a =$ the optimistic completion time estimate, $b =$ the pessimistic completion time estimate, and $m =$ the most likely completion time estimate.

In the three time estimate approach, the critical path is determined as if the mean times for the activities were fixed times. The overall project completion time is assumed to have a normal distribution with mean equal to the sum of the means along the critical path and variance equal to the sum of the variances along the critical path.

In the CPM (Critical Path Method) approach to project scheduling, it is assumed that the normal time to complete an activity, $j$, which can be met at a normal cost, $C_j$, can be crashed to a reduced time, $y$, under maximum crashing for an increased cost, $C_j$.

**PERT/CPM Critical Path Procedure**

Step 1: Develop a list of the activities that make up the project.

Step 2: Determine the immediate predecessors for each activity in the project.

Step 3: Estimate the completion time for each activity.

Step 4: Draw a project network depicting the activities (or nodes) and immediate predecessors listed in Steps 1 and 2.

Step 5: Use the project network and the activity time estimates to determine the earliest start and the earliest finish time for each activity by making a forward pass through the network. The earliest finish time for the last activity in the project identifies the total time required to complete the project.

Step 6: Use the project completion time identified in Step 5 as the latest finish time for the last activity, and make a backward pass through the network to identify the latest start and
Step 7: Use the difference between the latest start time and the earliest start time for each activity to determine the slack for the activity.

Step 8: Find the activities with zero slack; these are the critical activities.

Step 9: Use the information from Steps 5 and 6 to develop the activity schedule for the project.

Before constructing the PERT network, summarize in the following table the immediate predecessor activities for each activity.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Immediate Predecessors</th>
<th>Completion Times</th>
<th>Hr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>A</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>C</td>
<td>A</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>O</td>
<td>B</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>E</td>
<td>C</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>F</td>
<td>B, C</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>O, E</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>H</td>
<td>C</td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

To construct a PERT network, there must be a node for each activity. An activity's immediate predecessors are the nodes that immediately precede it. This is a network flow diagram that would support overlapping/rolling starts of activities.
3. Inventory Modeling

This is not a standalone document. It is coordinated with your *Quantitative Analysis for Management* text. Render, Stair, & Hanna (2012, pp. 195-215) note:

Inventory control is important to most organizations, namely the high cost of carrying too much inventory and the problems of stockouts (lost customers and reduced market share as a result of too little inventory).

There are two fundamental decisions in controlling inventory: (1) how much to order and (2) when to order. The objective is to minimize total inventory cost. Common inventory costs are:

- Cost of the items (purchase or material cost)
- Cost of ordering
- Cost of carrying or holding inventory
- Cost of stockouts

The economic order quantity (EOQ) model is one of the oldest and commonly used inventory control techniques. It has the following important assumptions:

1. Demand is known and constant.
2. Lead time is known and constant.
3. Receipt of inventory is instantaneous.
4. Purchase cost per unit is constant throughout the year.
5. The only variable costs are the cost of placing an order, ordering cost, and the cost of holding or storing inventory over time, holding or carrying cost; these are constant throughout the year.
6. Orders are placed so that stockouts or shortages are avoided completely.

The equation for EOQ is:

\[
EOQ = Q^* = \sqrt{\frac{2DC_o}{C_h}}
\]

Where:

- \( Q \) = number of pieces to order
- EOQ = \( Q^* \) = optimal number of pieces to order
- \( D \) = annual demand in units for the inventory item
- \( C_o \) = ordering cost of each order
- \( C_h \) = holding or carrying cost per unit per year

**Example:** Paul Peterson is the inventory manager for Office Supplies, Inc., a large office supply warehouse. The annual demand for paper punches is 20,000 units. The ordering cost is $100 per order, and the carrying cost is $5 per unit per year. The following equation can be used to compute the economic order quantity.

\[
Q^* = \sqrt{\frac{2(20,000)(100)}{5}} = 894
\]

Quantity discounts are another important consideration. Many suppliers provide a lower cost for ordering in larger than your typical economic order quantity. To determine if you should take the quantity discount, you need to compare total annual costs of the alternatives.
Paul Peterson (see above example) has found a supplier of hole punches that offers quantity discounts. The annual demand is 20,000 units, the ordering cost is $100 per order, and the carrying cost is 0.5 of the unit price. For quantities that vary from 0 to 1,999, the unit price is $10. The price is $9.98 for quantities that vary from 2,000 units to 3,999 units, and $9.96 for quantities that vary from 4,000 to 10,000 units. Should Paul take the quantity discount?

To solve this problem, we begin by computing the economic order quantity. This is done using the same equation as above.

\[ Q^* = \sqrt{\frac{2 \times 20,000 \times 100}{10 \times 0.5}} = 894 \]

The table below shows the results of the total cost analysis. Note that the order quantity had to be adjusted to 2,000 and 4,000 units for the last two quantity discounts. In this case, the best decision is not to take the quantity discount. The order quantity is 894 units with a total cost of $204,472.

<table>
<thead>
<tr>
<th>Discount Number</th>
<th>Unit Price</th>
<th>Order Quantity</th>
<th>Material Cost</th>
<th>Ordering Cost</th>
<th>Carrying Cost</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$10.00</td>
<td>894</td>
<td>$200,000</td>
<td>$2,236</td>
<td>$2,236</td>
<td>$204,472</td>
</tr>
<tr>
<td>2</td>
<td>9.98</td>
<td>2,000</td>
<td>199,600</td>
<td>1,000</td>
<td>4,990</td>
<td>205,590</td>
</tr>
<tr>
<td>3</td>
<td>9.96</td>
<td>4,000</td>
<td>199,200</td>
<td>500</td>
<td>9,960</td>
<td>209,660</td>
</tr>
</tbody>
</table>

Another important process is the use of safety stock. The basic Reorder Point (ROP) equation is

\[ \text{ROP} = d \times L \]

\(d = \text{daily demand or average daily demand}\)

\(L = \text{order lead time or the number of working days it takes to deliver an order (or average lead time)}\)

(ROP is expressed in units of product on hand)

A safety stock quantity is added to the equation to accommodate the uncertain demand during lead time or the uncertain delivery time. Thus

\[ \text{ROP} = d \times L + SS \]

\(SS = \text{Safety stock}\)

Graphically, safety stock looks like this:

4. Statistical process control - See section A 10 above.

5. Special topics

Break-even point analysis
This is not a standalone document. It is coordinated with your Quantitative Analysis for Management text. Render, Stair, & Hanna (2012, pp. 7-8) note:

The break-even point is the number of units sold that will result in zero profit.
A break-even point analysis is a popular method to assess certain actions. For example, look at the following situation: You are selling a product for $10 and its variable cost is $7. The product's fixed costs are $30,000. The size of the market you are operating is $1,000,000, and you have 12 percent market share.

Are you making any money?

The formula for profit (net income) is:

\[
\text{Profit} = \text{Revenue} - \text{total variable costs} - \text{fixed costs}
\]

Right now, your sales in dollars are: 1,000,000 * 0.12 = $120,000
Since you are selling each product at $10, your sales in units are: 120,000 / 10 = 12,000 units
Your total variable cost is: 12,000 * 7 = $84,000
Your fixed cost is: $30,000
Your total cost is: $114,000
Your profit is: 120,000 - 84,000 - 30,000 = $6,000

Another way to determine the profit is to do a BEP analysis.

Your contribution per unit is the difference between the price you sell each unit and the variable cost of producing it.

\[
\text{Contribution per unit} = \text{Price} - \text{variable cost}
\]

Thus, it is: 10 - 7 = $3

The formula for the break-even point is:

\[
\text{BEP} = \frac{\text{fixed costs}}{\text{contribution per unit}}
\]

BEP = 30,000 / 3 = 10,000 units (or $100,000).
This means that you need to sell 10,000 units (or to generate sales of $100,000) to break even. Since you are selling 12,000 units, you are making money. The difference is the contribution of the additional units beyond the BEP: (12,000 - 10,000) * 3 = $6,000. These additional dollars are your profit.

Now, let us assume that someone would suggest that spending an additional $30,000 in advertising would increase your market share to 20%. Should you go for it?

To answer this question, you will have to look at the marginal benefits and costs (marginal means the additional benefits and costs).

Additional benefits: sales will increase by $80,000 (at 20% market share your sales will grow to $200,000, which is an increase of 80,000 over the 120,000).
Additional contribution (marginal benefits): 80,000 / 10 = 8,000 units at $3 per unit = $24,000.
Additional cost (marginal cost): $30,000 for the increase in advertising.
Thus, the net effect is a loss of $6,000. The answer is no.

Another way to compute this is to look at the BEP. How many additional units will we have to sell to cover the new fixed cost of $30,000?

BEP = 30,000 / 3 = 10,000 units. The new advertising will increase our sales by 8,000 units (80,000 / 10 = 8,000), which is below the BEP. Thus, the proposal will have to be rejected.

Now, let’s say that the company wants to generate a profit of $9,000. How many units do you have to sell?
We can use the BEP formula. Instead of asking how many units we need to break even (i.e., the number of units to cover the fixed and variable costs), we ask for the number of units to cover the fixed costs, the variable costs, plus the profit we want to generate. The formula is:

\[
\text{Number of units} = \frac{(\text{fixed costs} + \text{profit})}{(\text{contribution per unit})}
\]

In our example, to generate profits of $9,000 we must sell the following number of units:

Number of units = \(\frac{30,000 + 9,000}{3} = 13,000\)

Many times you will be asked to calculate the BEP not in units but in dollars. If you already have it in units, then you just have to multiply the units by the price per unit. Thus, in the above example:

BEP = 30,000/3 = 10,000 units.

BEP (in $) = 10,000*10 = $100,000.

To break even, you have to sell 10,000 units; or to put it in another way, to generate revenue (sales) of $100,000.

An alternative method is to calculate the BEP in dollars directly. For that, we need to introduce the concept of contribution margin. The contribution margin is a relative measure of the contribution per unit. The frame of reference is the price per unit. Thus, the formula for the contribution margin is:

\[
\text{Contribution margin} = \frac{(\text{contribution per unit})}{(\text{price})}
\]

In our example,

Contribution margin (CM) = \(\frac{3}{10} = 0.3 \text{ or } 30\%\).

In other words, for every dollar we generate in sales, 30 cents are contributed to our bottom line.

The formula for BEP in dollars is:

\[
\text{BEP (in $)} = \frac{(\text{fixed costs})}{(\text{contribution margin})}
\]

In our example, the BEP in dollars is:

BEP (in $) = 30,000/0.3 = $100,000. Similar to the previous example, we can ask how many dollars in revenue we need to generate in order to end up with $9,000 in profits. The formula is:

\[
\text{Necessary revenue} = \frac{(\text{fixed costs} + \text{profit})}{(\text{contribution margin})}
\]

In our case,

Necessary revenue = \(\frac{30,000 + 9,000}{0.3} = $130,000\)

**Queuing Theory**

This is not a standalone document. It is coordinated with your *Quantitative Analysis for Management* text. Render, Stair, & Hanna (2012, pp. 26-28) note:

Queuing theory is the study of waiting lines (waiting for something to happen). Four characteristics of a queuing system are: (1) the manner in which customers arrive; (2) the time required for service; (3) the priority determining the order of service; and (4) the number and configuration of servers in the system.

In general, the arrival of customers into the system is a random event. Frequently the *arrival pattern* is
modeled as a Poisson process. The Poisson distribution defines the probability of \( x \) arrivals during a specified period of time.

A three part code is used to describe various queuing systems. Frequently used symbols for the arrival and service processes are: \( M \) - Markov distributions (Poisson/exponential), \( D \) - Deterministic (constant), and \( G \) - General distribution (with a known mean and variance).

Thus, the notation \( MIMIk \) refers to a queuing situation in which arrivals occur according to a Poisson distribution, service times follow an exponential distribution, and there are \( k \) servers each working at an identical service rate.

The arrivals, also called the calling population, are in two categories: infinite and finite. Finite is usually reserved for 30 or fewer in the calling population because the calculations are more difficult.

**Balking** refers to customers who refuse to join the queue. (Example: at a restaurant with a salad bar, people wait until the line is gone or only one to two people.) **Reneging** customers enter the queue but become impatient and leave without receiving their service. (Example: leaving a restaurant because the wait is more than an hour.)

The following is \( M/M/1 \) with Poisson arrivals and exponential service times and infinite population. We typically calculate the:

- Waiting time in the queue
- Waiting time in the system (includes waiting time in the queue)
- Number of units (such as people) in the queue
- Number of units in the system (includes the number in the queue)
- Percent idle time
- Percent of time more than \( n \) units are in the line

\( \lambda = \) mean number of arrivals per time period
\( \mu = \) mean number of customers or units served per time period

The arrival rate and service rate must be defined for the same time period.

**Average customers in the system** \( L \)

\[ L = \frac{\lambda}{\mu - \lambda} \]

**Average customers in the queue** \( L_q \)

\[ L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} \]

**Average time a customer spends in the system**

\[ W = \frac{1}{\mu - \lambda} \]

**Average time a customer spends in the queue** \( W_q \)

\[ W_q = \frac{\lambda}{\mu(\mu - \lambda)} \]
The utilization factor for the system $p$, the probability the service facility is being used:

$$\rho = \frac{\lambda}{\mu}$$

The probability that the number of customers in the system is greater than $k$, $P_{n>k}$:

$$P_{n>k} = \left(\frac{\lambda}{\mu}\right)^{k+1}$$

Example: Arnold’s Muffler Shop
Arnold’s mechanic can install mufflers at a rate of 3 per hour. $\mu = 3/hr$
Customers arrive at the rate of 2 per hour. $\lambda = 2/hr$

$L = \frac{\lambda}{\mu - \lambda} = \frac{2}{3-2} = \frac{2}{1} = 2$ cars in the system

$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{2^2}{3(3-2)} = \frac{4}{3(1)} = 1.33$ cars in the queue

$W = \frac{1}{\mu - \lambda} = \frac{1}{3-2} = 1$ hour wait time

$W_q = \frac{\lambda}{\mu(\mu - \lambda)} = \frac{2}{3} \text{ hour} = .67$ hour wait in the queue

$$\rho = \frac{\lambda}{\mu} = \frac{2}{3} = 0.67 = \text{mechanic is busy 67% of the time}$$

<table>
<thead>
<tr>
<th>$K$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.677</td>
</tr>
<tr>
<td>1</td>
<td>0.444</td>
</tr>
<tr>
<td>2</td>
<td>0.296</td>
</tr>
<tr>
<td>3</td>
<td>0.198</td>
</tr>
<tr>
<td>4</td>
<td>0.132</td>
</tr>
<tr>
<td>5</td>
<td>0.088</td>
</tr>
<tr>
<td>6</td>
<td>0.058</td>
</tr>
<tr>
<td>7</td>
<td>0.039</td>
</tr>
</tbody>
</table>

**Decision Analysis**
This is not a standalone document. It is coordinated with your *Quantitative Analysis for Management* text. Render, Stair, & Hanna (2012, pp. 26-28) note:

As you may remember, a decision problem is characterized by decision alternatives, states of nature, and the resulting payoffs.
The decision alternatives are the different strategies the decision maker can employ. The states of nature refer to future events, not under the control of the decision maker, which may occur. States of nature should be defined so that they are mutually exclusive and collectively exhaustive.

For each decision of an alternative and state of nature, there is a resulting payoff. These are often represented in matrix form called a payoff table. A decision is said to dominate another decision if the payoffs for every state of nature for one is at least equal to the corresponding payoffs for the other and is greater for at least one state of nature.

If the decision maker does not know with certainty which state of nature will occur, then he or she is said to be engaging in decision making under uncertainty.

Three commonly used criteria for decision making under uncertainty when probability information regarding the likelihood of the states of nature is unavailable are:

1. Optimistic approach (maximax)
2. Conservative approach (minimin)
3. Criterion of realism
4. Equally likely
5. Minimax regret

The optimistic approach would be used by an optimistic decision maker. The decision with the largest possible payoff is chosen. (If the payoff table were in terms of costs, the decision that had the lowest cost would be chosen.)

Note: The text below is also not a standalone document. It is coordinated with your text An Introduction to Management Science. Anderson, Sweeney, & Williams (2008, pp. 650-651) state:

The conservative approach would be used by a conservative decision maker. For each decision, the minimum payoff is listed. Then the decision corresponding to the maximum of these minimum payoffs is selected. Choose the maximum of the minimums in the weakest column.

The mini max regret approach requires the construction of a regret table or an opportunity loss table. This is done by calculating for each state of nature the difference between each payoff and the largest payoff for that state of nature. Then, using this regret table, the maximum regret for each possible decision is listed. The decision chosen is the one corresponding to the minimum of the maximum regrets.

An equally likely approach involves finding the average payoff for each alternative and selecting the alternative with the highest average. This approach assumes that all probabilities of occurrence for the states of nature are equal, and thus each state of nature is equally likely.

Criterion of realism approach is a compromise between an optimistic and a pessimistic decision. A coefficient of realism is selected. This measures the degree of optimism of the decision maker. The coefficient is between 0 and 1. When the coefficient is 1, the decision maker is 100% optimistic; when the coefficient is 0, the decision maker is 100% pessimistic about the future.

If probabilistic information regarding the states of nature is available, use the expected value (EV) approach. Here the expected return for each decision is calculated by summing the products of the payoff under each state of nature and the probability of the respective state of nature occurring. The decision yielding the best expected return is chosen.

Risk analysis is the study of the possible payoffs associated with a decision alternative or a decision strategy. A tool used in risk analysis is graphically depicting the risk profile for a decision alternative. The risk profile shows the possible payoffs for the decision alternative along with their associated probabilities. Sometimes a review of the risk will cause the decision maker to choose another decision alternative even though the expected value of the other alternative is not as good.
Frequently, information is available which can improve the probability estimates for the states of nature. The expected value of perfect information (EVPI) is the increase in the expected profit that if one knew, with certainty, which state of nature would occur. The EVPI can be calculated as follows:

1. determine the optimal return corresponding to each state of nature;
2. compute the expected value of these optimal returns;
3. subtract the EV of the optimal decision from the amount determined in step (2).

References


