Project Management

The chapter in this section covers various topics related to managing projects.

CHAPTER 17

Project Management

LEARNING OBJECTIVES

After completing this chapter, you should be able to:

1. Discuss the behavioral aspects of projects in terms of project personnel and the project manager.
2. Discuss the nature and importance of a work breakdown structure in project management.
3. Give a general description of PERT/CPM techniques.
4. Construct simple network diagrams.
5. List the kinds of information that a PERT or CPM analysis can provide.
6. Analyze networks with deterministic times.
7. Analyze networks with probabilistic times.
8. Describe activity "crashing" and solve typical problems.

CHAPTER OUTLINE

The International Space Station Project (ISS)

www.nasa.com

Hopefully, by 2006 the space station assembly will be complete and functional 220 miles above the earth at a cost of more than $40 billion. The station will cover an area of approximately two football fields and have a housing unit the size of a jumbo jet. This space station will conduct research concerning human life and production of products in space. The complexity of the space station and the coordination required among 16 partner nations is daunting.

Russia launched the first piece of the space station on November 20, 1998. This module provides power for future assembly operations. On December 3 of the same year, the United States launched the first space component hub, which serves as the connector for other pieces of the station. This hub also serves as housing for up to seven astronauts. In total it is planned that 34 space deliveries will transport more than

900,000 pounds of materials needed to assemble the completed station. It is estimated that 144 to 160 two-person space walks and 1,800 hours of construction will be needed to assemble the station.

The project has had a bumpy ride to date, and more problems are expected before the project is complete. The project is billions of dollars over budget; the total cost could reach $96 billion. The collapse of the Russian economy has caused delays on their development of a major module of the station. Financial support has been needed to get the Russian effort back on track. Redesign has been continual because of the multiple needs of the partners of the 16 nations; these changes have resulted in costly delays. Some of the risks have been underestimated, and reassessment has suggested failures along the way are inevitable—delivery vehicles blowing up, docking problems, even someone killed during construction. There have been calls to "pull the plug" on the project. Nevertheless, the project moves forward.

(continued)

(concluded)

The ESA (European Space Agency) is coordinating several ISS projects linked to the development of an ATV (automatic transfer vehicle—not related to all-terrain vehicles), which will supply nine tons of cargo and fuel to the space station. These projects, which include countries and companies from France, Germany, Russia, and Italy, will take place in their respective countries under the guidance of ESA. DaimlerChrysler will produce the estimated 15 ATVs scheduled for delivery between 2003 and 2013. Standard penalty clauses for late delivery are in place. Incentive provisions for continuous improvement of mis-

sion success and payload mass are included in the multination endeavor.

Problems are deemed surmountable, and plans to complete the assembly by 2006 are in place. It is expected that the world-class laboratory will lead to discoveries that will touch the lives of everyone in the world. Perhaps the lessons learned will be as valuable as the discoveries.

The space station is a huge project. Here are some examples of large projects. The Olympic Games—summer games, winter games, you name it—involve much more than the festivities, the excitement, national pride, and competition among athletes. They all involve a tremendous amount of planning, preparation, and coordinating work that needs to get done before and during the games. Athletes' living quarters and training facilities must be provided, competition schedules must be developed, arrangements for televising events must be made, equipment and crews must be coordinated, transportation and hotel accommodations must be made, and many other activities that go on behind the scenes must be planned and managed so that everything goes off smoothly.

The Microsoft Corporation periodically releases new or updated software. Each release is the result of many people working countless hours writing code, testing programs, and revising code. Design, production, and marketing efforts also have to be coordinated. The reputation and profits of the company are closely related to successful software development.

www.microsoft.com

A few years ago major portions of the Santa Ana freeway in California were destroyed by an earthquake and had to be rebuilt. Thousands of activities had to be planned and coordinated. Thousands of pieces of construction equipment and tons of material had to be coordinated. The rebuilding was successful, and actually completed 70 days ahead of schedule.

The movie *Titanic,* released in late 1997, was the leading box office money maker of all time. And a good thing too, because over $250 million was spent on making the movie, also a record. Imagine the immense job of planning and coordinating the script writers, set designers, set builders, camera crews, directors, actors, makeup, costumes, advertising, and more that went into making that movie.

These are just a few examples of projects. In fact, there are projects going on all around us, every day. Some are large, some are small. Obviously, projects are an important way that certain kinds of work get done.

INTRODUCTION

projects Unique, one-time operations designed to accomplish a specific set of objectives in a limited time frame.

Managers typically oversee a variety of operations. Some of these involve routine, repetitive activities, but others involve *nonroutine* activities. Under the latter heading are **projects:** unique, one-time operations designed to accomplish a set of objectives in a limited time frame. Other examples of projects include constructing a shopping complex, merging two companies, putting on a play, and designing and running a political campaign. Examples of projects within business organizations include designing new products or services, designing advertising campaigns, designing information systems, reengineering a process, designing databases, software development, and designing web pages.

Projects may involve considerable cost. Some have a long time horizon, and some involve a large number of activities that must be carefully planned and coordinated. Most are expected

to be completed based on time, cost, and performance targets. To accomplish this, goals must be established and priorities set. Tasks must be identified and time estimates made. Resource requirements also must be projected and budgets prepared. Once under way, progress must be monitored to assure that project goals and objectives will be achieved.

The project approach enables an organization to focus attention and concentrate efforts on accomplishing a narrow set of performance objectives within a limited time and budget framework. This can produce significant benefits compared with other approaches that might be considered. Even so, projects present managers with a host of problems that differ in many respects from those encountered with more routine activities. The problems of planning and co-ordinating project activities can be formidable for large projects, which typically have thousands of activities that must be carefully planned and monitored if the project is to proceed according to schedule and at a reasonable cost.

The chapter introduces the basic concepts of project management. It begins with a brief discussion of some behavioral aspects of project management, along with some of the difficulties project managers are apt to encounter. The main portion of the chapter is devoted to a description of graphical and computational methods that are used for planning and scheduling projects.

Table 17.1 provides an overview of project management.

BEHAVIORAL ASPECTS OF PROJECT MANAGEMENT

Project management differs from management of more traditional activities mainly because of its limited time framework and the unique set of activities involved, which gives rise to a host of unique problems. This section describes more fully the nature of projects and their behavioral implications. Special attention is given to the role of the project manager.

The Nature of Projects

Projects go through a series of stages—a life cycle—which include project definition planning, execution of major activities, and project phaseout. During this life cycle, a variety of skill requirements are involved. The circumstances are analogous to constructing a house. Initially an idea is presented and its feasibility is assessed, then plans must be drawn up and approved by the owner and possibly a town building commission or other regulatory agency. Then a succession of activities occurs, each with its own skill requirements, starting with the site preparation, then laying the foundation, erecting the frame, roofing, constructing exterior walls, wiring and plumbing, installing kitchen and bathroom fixtures and appliances, interior finishing work, and painting and carpeting work. Similar sequences occur on large construction projects, in R&D work, in the aerospace industry, and in virtually every other instance where projects are being carried out.

Projects typically bring together people with diverse knowledge and skills, most of whom remain associated with the project for less than its full life. Some people go from project to project as their contributions become needed, and others are "on loan," either on a full-time or part-time basis, from their regular jobs. The latter is usually the case when a special project exists within the framework of a more traditional organization. However, some organizations are involved with projects on a regular basis; examples include consulting firms, architects, writers and publishers, and construction firms. In those organizations, it is not uncommon for some individuals to spend virtually all of their time on projects.

Some organizations use a *matrix organization* that allows them to integrate the activities of a variety of specialists within a functional framework. For instance, they have certain people who prepare proposals, others who concentrate exclusively on engineering, others who devote their efforts to marketing, and so on.

Fireworks over Rice-Eccles Olympic Stadium at the opening ceremony for the 2002 Winter Olympic games in Salt Lake City, Utah. Planning and managing the execution of the Olympic Games is a large project that includes long-term construction and improvement of sites as well as precisely timed activities such as the opening ceremony.

TABLE 17.1
Overview of project management

> **What is project management?** A team-based approach for managing projects.
> **How is it different from general operations management?**
> 1. Limited time frame
> 2. Narrow focus, specific objectives
> 3. Less bureaucratic
> **Why is it used?**
> 1. Special needs that don't lend themselves to functional management
> 2. Pressures for new or improved products or services, cost reduction
> **What are the key metrics?**
> 1. Time
> 2. Cost
> 3. Performance objectives
> **What are the key success factors?**
> 1. Top-down commitment
> 2. A respected and capable project manager
> 3. Enough time to plan
> 4. Careful tracking and control
> 5. Good communications
> **What are the major administrative issues?**
> 1. Executive responsibilities
>   a. Project selection
>   b. Selection of a project manager
>   c. Organizational structure (to whom will the project manager report?)
> 2. Organizational alternatives
>   a. Manage within functional unit
>   b. Assign a coordinator
>   c. Use a matrix organization with a project leader
> **What are the main tools?**
> 1. *Work breakdown structure.* An initial planning tool that is needed to develop a list of activities, activity sequences, and a realistic budget.
> 2. *Network diagram.* A "big picture" visual aid that is used to estimate project duration, identify activities that are critical for timely project completion, identify areas where slack time exists, and develop activity schedules.
> 3. *Gantt charts.* A visual aid used to plan and monitor individual activities.
> 4. *Risk management.* Analyses of potential failures or problems, assessment of their likelihood and consequences, and contingency plans.

Key Decisions in Project Management

Much of the success of projects depends on key managerial decisions over a sequence of steps:

* Deciding which projects to implement.
* Selecting the project manager.
* Selecting the project team.
* Planning and designing the project.
* Managing and controlling project resources.
* Deciding if and when a project should be terminated.

Deciding Which Projects to Implement This involves determining the criteria that will be used to decide which projects to pursue. Typical factors include budget, availability of appropriate knowledge and skill personnel, and cost–benefit considerations. Of course, other factors may override these criteria, factors such as availability of funds, safety issues, government-mandated actions, and so on.
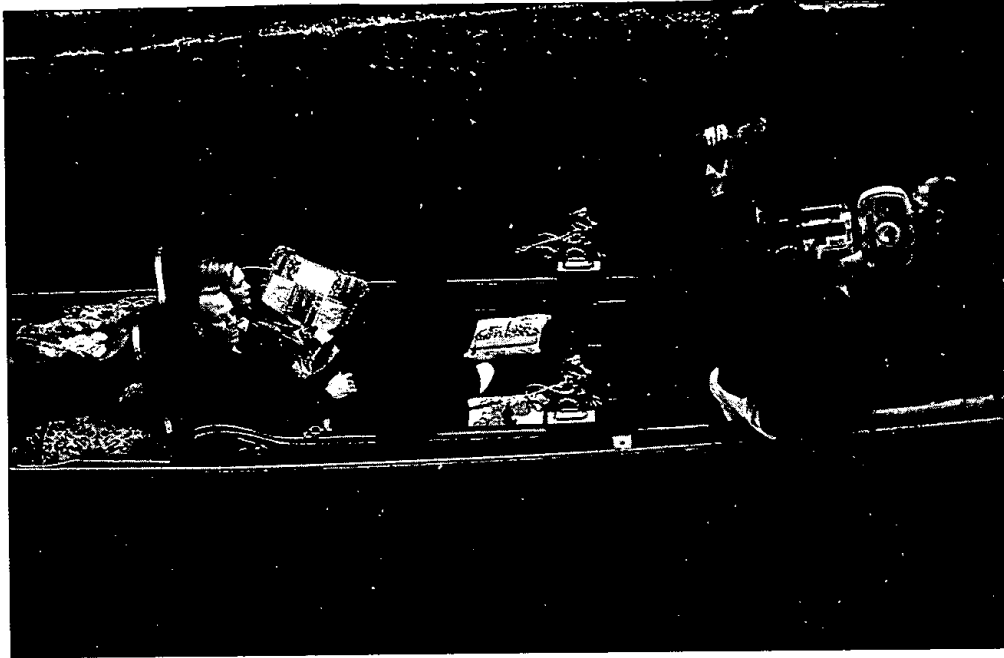
Making a film is a project involving people with a wide variety of talent and skills and may include scheduling multiple locations. This is Bruce Willis and Michelle Pfeifer on the set of "The Story of Us" in Venice.

Selecting the Project Manager The project manager is the central person in the project. The following section on project managers discusses this topic.

Selecting the Project Team The team can greatly influence the ultimate success or failure of a project. Important considerations include not only a person's knowledge and skill base, but also how well the person works with others (particularly those who have already been chosen for the project), enthusiasm for the project, other projects the person is involved in, and how likely those other projects might be to interfere with work on this project.

Planning and Designing the Project Project planning and design require decisions on project performance goals, a timetable for project completion, the scope of the project, what work needs to be done, how it will be done, if some portions will be outsourced, what resources will be needed, a budget, and when and how long resources will be needed.

Managing and Controlling Project Resources This involves managing personnel, equipment, and the budget; establishing appropriate metrics for evaluating the project; monitoring progress; and taking corrective action when needed. Also necessary are designing an information system and deciding what project documents should be generated, their contents and format, when and by whom they will be needed, and how often they should be updated.

Deciding If and When a Project Should Be Terminated Sometimes it is better to terminate a project than to invest any more resources. Important considerations here are the likelihood of success, termination costs, and whether resources could be better used elsewhere.

The Project Manager

The project manager bears the ultimate responsibility for the success or failure of the project. He or she must be capable of working through others to accomplish the objectives of the project. The project manager is responsible for effectively managing each of the following:

1. The *work*, so that all of the necessary activities are accomplished in the desired sequence, and performance goals are met.

2. The *human resources*, so that those working on the project have direction and motivation.

3. *Communications*, so that everybody has the information they need to do their work.

4.   *Quality*, so that performance objectives are realized.

5.   *Time*, so that the project is completed on schedule.

6.   *Costs*, so that the project is completed within budget.

To effectively manage a project, a project manager must employ a certain set of skills. The skills include the ability to motivate and direct team members; make trade-off decisions; expedite the work when necessary; put out fires; and monitor time, budget, and technical details. For projects that involve fairly well-defined work, those skills will often suffice. However, for projects that are less well defined, and thus have a higher degree of uncertainty, the project manager also must employ strong leadership skills. These include the ability to adapt to changing circumstances that may involve changes to project goals, technical requirements, and project team composition. As a leader, the project manager not only must be able to deal with these issues; he or she also must recognize the need for change, decide what changes are necessary, and work to accomplish them.

The job of project manager can be both difficult and rewarding. The manager must coordinate and motivate people who sometimes owe their allegiance to other managers in their functional areas. In addition, the people who work on a project frequently possess specialized knowledge and skills that the project manager lacks. Nevertheless, the manager is expected to guide and evaluate their efforts. Project managers often must function in an environment that is beset with uncertainties. Even so, budgets and time constraints are usually imposed, which can create additional pressures on project personnel. Finally, the project manager may not have the authority needed to accomplish all the objectives of the project. Instead, the manager sometimes must rely on persuasion and the cooperation of others to realize project goals.

Ethical issues often arise in connection with projects. Examples include the temptation to understate costs or to withhold information in order to get a project approved, pressure to alter or make misleading statements on status reports, falsifying records, compromising workers' safety, and approving substandard work. It is the responsibility of managers at all levels to maintain and enforce ethical standards. Moreover, employees often take their cue from managers' behavior, so it is doubly important for managers to model ethical behavior. The Project Management Institute (PMI) has a website (www.pmi.org) that includes a code of ethics for project managers, in addition to other useful information about project management.

The position of project manager has high visibility. The rewards of the job of project manager come from the creative challenges of the job, the benefits of being associated with a successful project (including promotion and monetary compensation), and the personal satisfaction of seeing it through to its conclusion.

Project Champions

project champion A person who promotes and supports a project.

Some companies make use of **project champions**. These are people, usually within the company, who promote and support the project. They can be instrumental in facilitating the work of the project manager by "talking up" the project to other managers who might be asked to share resources with the project team as well as employees who might be asked to work on parts of the project. The work that a project champion does can be critical to the success of a project, so it is important for team members to encourage and communicate with the project champion.

Project Managers Have Never Been More Critical

READING

Bob Weinstein

Not many people can eloquently describe the changing role of project managers as well as Jonathan Gispan, who teaches the skill to managers at the Lockheed Martin Corporation division in

King of Prussia, PA. Gispan spent 38 years of his career working at GE Aerospace as a project manager.

(concluded)

Project managers have never been more critical, according to this veteran techie. The function's core responsibilities haven't changed, but the players and playing field are very different than they were five years ago.

Gispan views the project manager as the focal point of a complex relationship matrix that includes customers, workers, vendors and bosses. "Project managers make things happen," he says. "Not only do they manage a project from start to finish, they also manage the people who make it happen."

The job requires a jack-of-all-trades who can both understand a problem and solve it. "Project managers must successfully manage the total lifecycle of a project," says Gispan. "The job requires a super-organized person who knows how to get a project completed quickly and inexpensively and boasts the advanced communication skills necessary to work closely with vendors, project teams and senior management."

Project managers have always had to deal with changing relationships. The big change, according to Gispan, is the speed of change, which has made the project manager's job more difficult. "The changes taking place now are faster and more dramatic than in the past," he says. "Customers are less inclined to give you money for development work, for example. They'd rather save money by using off-the-shelf capabilities. Budgets that used to be reasonably substantial and healthy are now tight."

Gispan adds that corporate mergers and consolidations have complicated the picture. Project managers must deal with complex, often labyrinthine decision-making structures. Getting quick decisions requires work and persistence.

Further complicating the situation is a high worker attrition rate. "A decade ago, the attrition rate was about 3 percent; today it has jumped to almost 10 percent," Gispan explains. "It means project managers have to constantly put together new teams."

Through it all, the project manager must be clear-headed and keep tight rein on projects. It sounds intimidating, yet Gispan sees it as a positive. "The current marketplace keeps you on your toes," he says. "Even though everything is changing at what seems like an overwhelming pace, it is best to view change as a motivator that allows you to keep pace with tech-nology and get better at your job. The idea is not merely to cope with change but to thrive in it. It is more of an attitude than anything else."

What does it take to be a project manager? "You need a good technical base," says Karen Nichols, project manager at EWP Engineering, Inc., a consulting engineering company in Salt Lake City, UT. "That translates to at least five to seven years in the trenches."

That's for starters. "But you also need to know how the business side of the equation works," says Nichols. "This is tough for some technical people because it requires the ability to understand how the two different sides of the business mesh. Project managers must also be familiar with accounting methods as well as sales and marketing strategies. In short, they must know how to manage a project so it makes a profit."

If you're not sure if you're up to the job, Nichols suggests testing the waters by watching other project managers at work. "You're more likely to get this opportunity at a mid-size to large company," she says.

If you think you have what it takes, ask to work with a project manager. "It won't take you long to find out whether you have an aptitude and natural feel for the work," says Nichols. "Not everyone is happy—or capable of—juggling many balls in the air at once."

Once you get your feet wet, you can move up quickly, adds Gispan. "After you've proven to management that you're organized and can manage several tasks, it won't be long before you are managing everything associated with a program. Get good at it and you'll move up the ranks from project manager to general manager where you are running several company businesses."

And, don't be confused by job titles. Each company has its own unique title for the project manager role. Whatever the title, a project manager is easy to spot. "It's the person who negotiates with all the key players and makes things happen," says Gispan. "Technology companies would perish without them."

Certification

The Project Management Institute (PMI) administers a globally recognized, examination-based professional certification program. There are several levels of certification: Associate, Project Management Professional, and Added Qualification. Candidates for the Associate and Professional levels must meet specific education and experience requirements and agree to adhere to a code of professional conduct. An advanced level of certification, Certificate of Added Qualification, is designed for those who hold the Professional certification who want to demonstrate a higher level of project management specialization within a specific industry. Both the Professional certification and the Advanced certification must demonstrate an ongoing professional commitment to the field of project management by satisfying PMI's Continuing Certification Requirements Program.[1]

[1] www.pmi.org.

The Pros and Cons of Working on Projects

People are selected to work on special projects because the knowledge or abilities they possess are needed. In some instances, however, their supervisors may be reluctant to allow them to interrupt their regular jobs, even on a part-time basis, because it may require training a new person to do a job that will be temporary. Moreover, managers don't want to lose the output of good workers. The workers themselves are not always eager to participate in projects because it may mean working for two bosses who impose differing demands, it may disrupt friendships and daily routines, and it presents the risk of being replaced on the current job. Furthermore, there may be fear of being associated with an unsuccessful project because of the adverse effect it might have on career advancement. In too many instances, when a major project is phased out and the project team disbanded, team members tend to drift away from the organization for lack of a new project and the difficulty of returning to former jobs. This tendency is more pronounced after lengthy projects and is less likely to occur when a team member works on a part-time basis.

In spite of the potential risks, people are attracted by the potential rewards of being involved in a project. One is the dynamic environment that surrounds a project, often a marked contrast to the staid environment of a routine in which some may feel trapped. Some individuals seem to thrive in more dynamic environments; they welcome the challenge of working under pressure and solving new problems. Then, too, projects may present opportunities to meet new people and to increase future job opportunities, especially if the project is successful. In addition, association with a project can be a source of status among fellow workers. Finally, working on projects frequently inspires a team spirit, increasing morale and motivation to achieve successful completion of project goals.

PROJECT LIFE CYCLE

The size, length, and scope of projects vary widely according to the nature and purpose of the project. Nevertheless, all projects have something in common: They go through a life cycle, which typically consists of four phases.

1. *Definition.* This has two parts: (*a*) *concept*, at which point the organization recognizes the need for a project or responds to a request for a proposal from a potential customer or client, and (*b*) *feasibility analysis*, which examines the expected costs, benefits, and risks of undertaking the project.

2. *Planning*, which spells out the details of the work and provides estimates of the necessary human resources, time, and cost.

3. *Execution*, during which the project itself is done. This phase often accounts for the majority of time and resources consumed by a project.

4. *Termination*, during which closure is achieved. Termination can involve reassigning personnel and dealing with any leftover materials, equipment (e.g., selling or transferring equipment), and any other resources associated with the project.

It should be noted that the phases can overlap, so that one phase may not be fully complete before the next phase begins. This can reduce the time necessary to move through the life cycle, perhaps generating some competitive advantage and cost saving. Although subsequent decisions in an earlier phase may result in waste for some portion of the activity in a following phase, careful coordination of activities can minimize that risk.

Figure 17.1 illustrates the phases in a project life cycle.

WORK BREAKDOWN STRUCTURE

Because large projects usually involve a very large number of activities, planners need some way to determine exactly what will need to be done so that they can realistically estimate how long it will take to complete the various elements of the project and how much it will cost. They
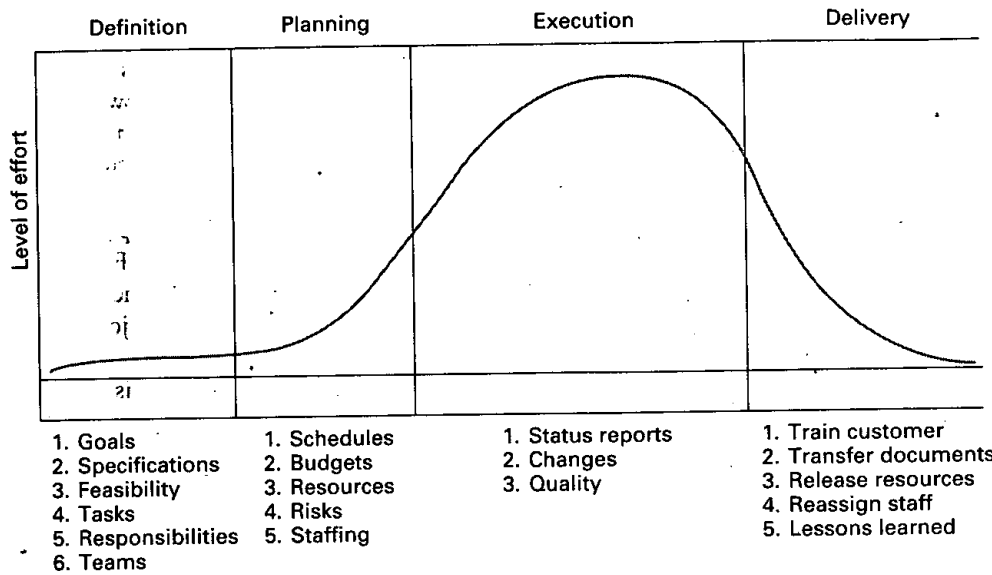
Definition | Planning | Execution | Delivery

Level of effort

1. Goals	1. Schedules	1. Status reports	1. Train customer
2. Specifications	2. Budgets	2. Changes	2. Transfer documents
3. Feasibility	3. Resources	3. Quality	3. Release resources
4. Tasks	4. Risks		4. Reassign staff
5. Responsibilities	5. Staffing		5. Lessons learned
6. Teams			

Level 1   Project

Level 2

Level 3

Level 4
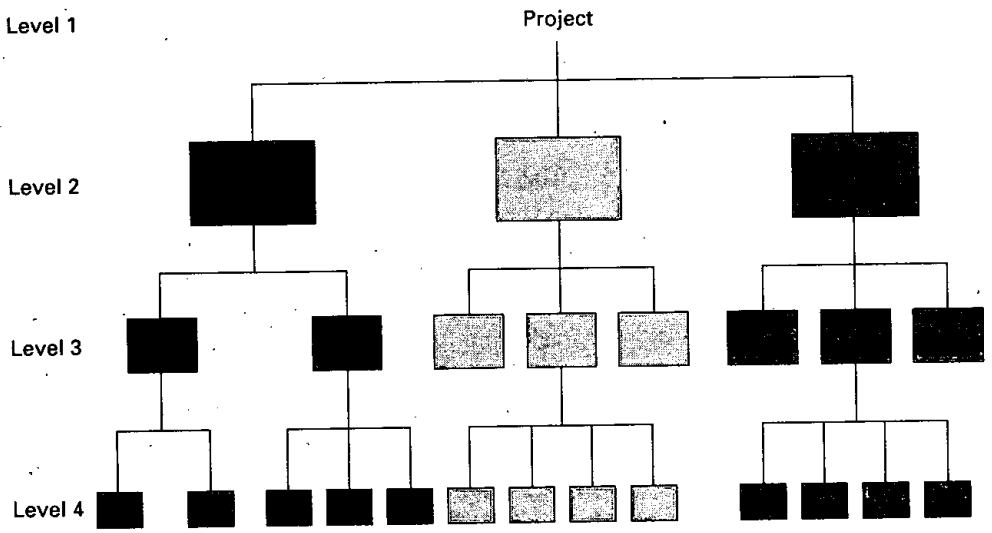
FIGURE 17.2

Schematic of a work breakdown structure

often accomplish this by developing a **work breakdown structure (WBS)**, which is a hierarchical listing of what must be done during the project. This methodology establishes a logical framework for identifying the required activities for the project (see Figure 17.2). The first step in developing the work breakdown structure is to identify the major elements of the project. These are the Level 2 boxes in Figure 17.2. The next step is to identify the major supporting activities for each of the major elements—the Level 3 boxes. Then, each major supporting activity is broken down into a list of the activities that will be needed to accomplish it—the Level 4 boxes. (For purposes of illustration, only a portion of the Level 4 boxes are shown.) Usually there are many activities in the Level 4 lists. Large projects involve additional levels, but Figure 17.2 gives you some idea of the concept of the work breakdown structure.

Developing a good work breakdown structure can require substantial time and effort due to the uncertainties associated with a project and/or the size of the project. Typically the portion of time spent on developing the work breakdown structure greatly exceeds the time spent on actually developing a project schedule. The importance of a work breakdown structure is underscored by the fact that the activity list that results serves as the focal point for planning and doing the project. Moreover, the work breakdown structure is the basis for developing time and cost estimates.

work breakdown structure (WBS) A hierarchical listing of what must be done during a project.

FIGURE 17.3

Gantt chart for bank example
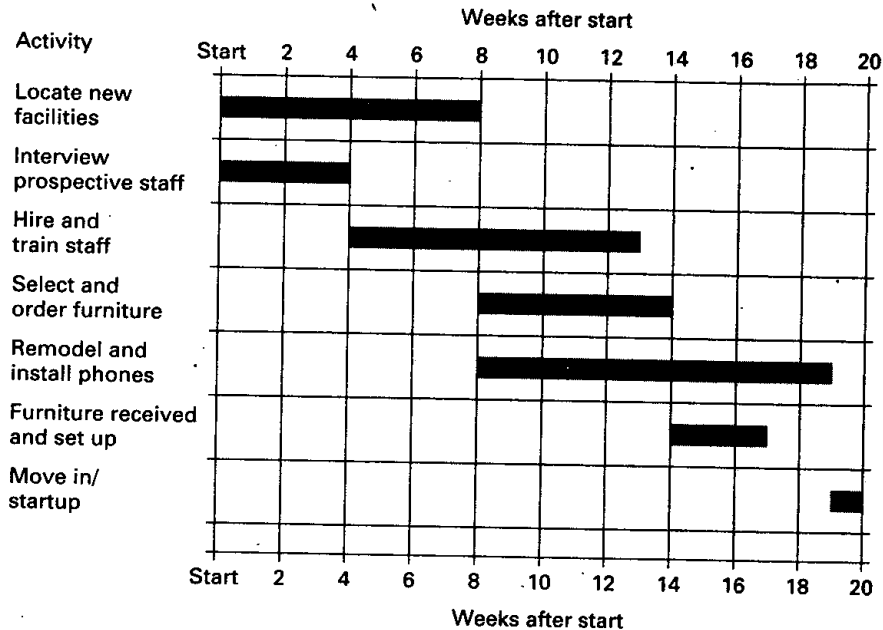
PLANNING AND SCHEDULING WITH GANTT CHARTS

The Gantt chart is a popular tool for planning and scheduling *simple* projects. It enables a manager to initially schedule project activities and then to monitor progress over time by comparing planned progress to actual progress. Figure 17.3 illustrates a Gantt chart for a bank's plan to establish a new direct marketing department. To prepare the chart, the vice president in charge of the project had to first identify the major activities that would be required. Next, time estimates for each activity were made, and the sequence of activities was determined. Once completed, the chart indicated which activities were to occur, their planned duration, and when they were to occur. Then, as the project progressed, the manager was able to see which activities were ahead of schedule and which were delaying the project. This enabled the manager to direct attention where it was needed most to speed up the project in order to finish on schedule.

Aside from being a visual tool, an advantage of a Gantt chart is its simplicity. However, Gantt charts fail to reveal certain relationships among activities that can be crucial to effective project management. For instance, if one of the early activities in a project suffers a delay, it would be important for the manager to be able to easily determine which later activities would result in a delay. Conversely, some activities may safely be delayed without affecting the overall project schedule. The Gantt chart does not directly reveal this. On more complex projects, it is often used in conjunction with a *network diagram*, the subject of the following section, for scheduling purposes.

PERT AND CPM

PERT Program evaluation and review technique, for planning and coordinating large projects.

CPM Critical path method, for planning and coordinating large projects.

PERT (program evaluation and review technique) and **CPM** (critical path method) are two of the most widely used techniques for planning and coordinating large-scale projects. By using PERT or CPM, managers are able to obtain

1. A graphical display of project activities.
2. An estimate of how long the project will take.
3. An indication of which activities are the most critical to timely project completion.
4. An indication of how long any activity can be delayed without delaying the project.

Although PERT and CPM were developed independently, they have a great deal in common. Moreover, many of the initial differences between them have disappeared as users borrowed certain features from one technique for use with the other. For practical purposes, the two techniques now are the same; the comments and procedures described will apply to CPM analysis as well as to PERT analysis of projects.

The Network Diagram

One of the main features of PERT and related techniques is their use of a **network or precedence diagram** to depict major project activities and their sequential relationships. There are two slightly different conventions for constructing these network diagrams. Under one convention, the *arrows* designate activities; under the other convention, the *nodes* designate activities. These conventions are referred to as **activity-on-arrow (AOA)** and **activity-on-node (AON)**. **Activities** consume resources and/or *time*. The nodes in the AOA approach represent the activities' starting and finishing points, which are called **events**. Events are points in time. Unlike activities, they consume neither resources nor time. The nodes in an AON diagram represent activities.

Both conventions are illustrated in Figure 17.4, using the bank example that was depicted in the Gantt chart in Figure 17.3. Compare the two. In the AOA diagram, the arrows represent activities and they show the sequence in which certain activities must be performed (e.g., Interview precedes Hire and train); in the AON diagram, the arrows show only the sequence in which certain activities must be performed while the nodes represent the activities. Activities in AOA networks can be referred to in either of two ways. One is by their endpoints (e.g., activity 2-4) and the other is by a letter assigned to an arrow (e.g., activity *c*). Both methods are illustrated in this chapter. Activities in AON networks are referred to by a letter (or number) assigned to a node. Although these two approaches are slightly different, they both show sequential relationships—something Gantt charts don't. Note that the AON diagram has a starting node, S, which is actually not an activity but is added in order to have a single starting node.

Despite these differences, the two conventions are remarkably similar, so you should not encounter much difficulty in understanding either one. In fact, there are convincing arguments for having some familiarity with *both* approaches. Perhaps the most compelling is that both approaches are widely used. However, any particular organization would typically use only one approach, and employees would have to work with that approach. Moreover, a contractor doing work for the organization may be using the other approach, so employees of the organization who deal with the contractor on project matters would benefit from knowledge of the other approach.

network (precedence) diagram Diagram of project activities that shows sequential relationships by use of arrows and nodes.

activity-on-arrow (AOA) Network diagram convention in which arrows designate activities.

activity-on-node (AON) Network diagram convention in which nodes designate activities.
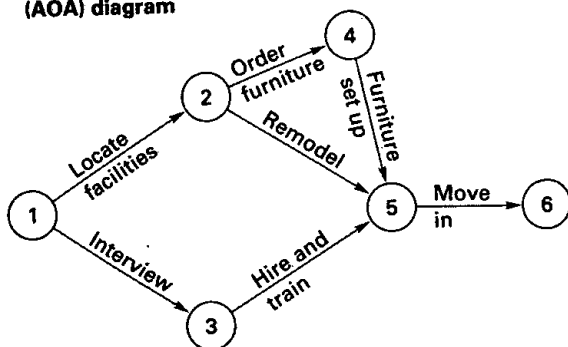
activities Project steps that consume resources and/or time.

events The starting and finishing of activities, designated by nodes in the AOA convention.
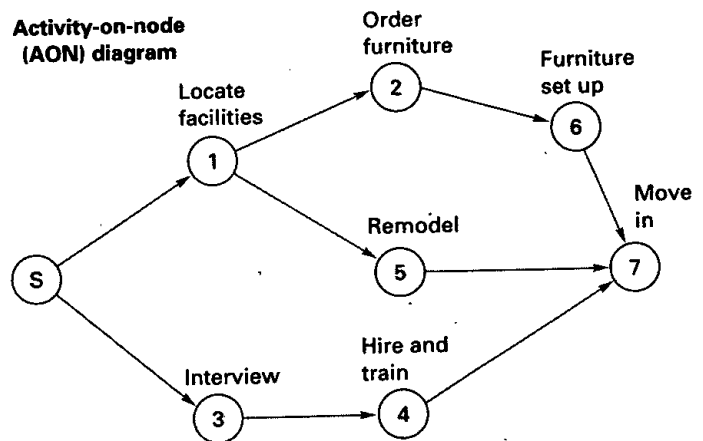
FIGURE 17.4

A simple project network diagram

path A sequence of activities that leads from the starting node to the finishing node.

Of particular interest to managers are the *paths* in a network diagram. A **path** is a sequence of activities that leads from the starting node to the ending node. For example, in the AOA diagram, the sequence 1-2-4-5-6 is a path. In the AON diagram, S-1-2-6-7 is a path. Note that in both diagrams there are three paths. One reason for the importance of paths is that they reveal *sequential relationships*. The importance of sequential relationships cannot be overstated: If one activity in a sequence is delayed (i.e., late) or done incorrectly, the start of all following activities on that path will be delayed.

Another important aspect of paths is the length of a path: How long will a particular sequence of activities take to complete? The length (of time) of any path can be determined by summing the expected times of the activities on that path. The path with the longest time is of particular interest because it governs project completion time. In other words, expected project duration equals the expected time of the longest path. Moreover, if there are any delays along the longest path, there will be corresponding delays in project completion time. Attempts to shorten project completion must focus on the longest sequence of activities. Because of its influence on project completion time, the longest path is referred to as the **critical path**, and its activities are referred to as **critical activities**.

critical path The longest path; determines expected project duration.

critical activities Activities on the critical path.

Paths that are shorter than the critical path can experience some delays and still not affect the overall project completion time as long as the ultimate path time does not exceed the length of the critical path. The allowable slippage for any path is called **slack**, and it reflects the difference between the length of a given path and the length of the critical path. The critical path, then, has zero slack time.

slack Allowable slippage for a path; the difference between the length of a path and the length of the critical path.

Network Conventions

Developing and interpreting network diagrams requires some familiarity with networking conventions. Table 17.2 illustrates some of the most basic and most common features of network diagrams. This will provide sufficient background for understanding the basic concepts associated with precedence diagrams and allow you to solve typical problems.

A special feature that is sometimes used in A-O-A networks to clarify relationships is a *dummy activity*. In order to recognize the need to use a dummy activity using the AOA approach when presented with a list of activities and the activities each precedes, examine the "Immediate Predecessor" list. Look for instances where multiple activities are listed, such as *a, b* in the following list. If *a* or *b* appears separately in the list (as *a* does here), a dummy will be needed to clarify the relationship (see the last diagram in Table 17.2).

Activity	Immediate Predecessor
a	—
b	—
c	a
d	a, b

Here are two more AOA conventions:

For reference purposes, nodes are numbered typically from left to right, with lower numbers assigned to preceding nodes and higher numbers to following nodes:
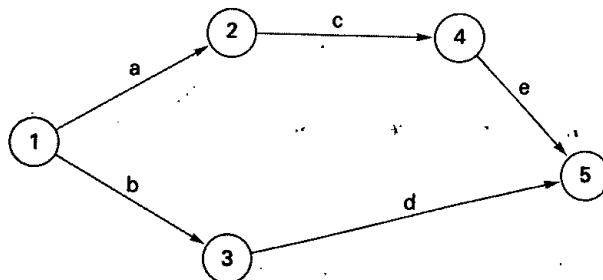
TABLE 17.2 Network conventions

AOA	INTERPRETATION	AON

Activities must be completed in sequence: first *a*, then *b*, and then *c*.

Both *a* and *b* must be completed before *c* can start.

Activity *a* must be completed before *b* or *c* can start.

Both *a* and *b* must be completed before *c* or *d* can start.

Use a dummy activity to clarify relationships:

1. To separate two activities that have the same starting and ending nodes. (No dummy needed)

2. When activities share some, but not all, precedence activities. (No dummy needed)

Dummy activity

Starting and ending arrows are sometimes used during development of a network for increased clarity:
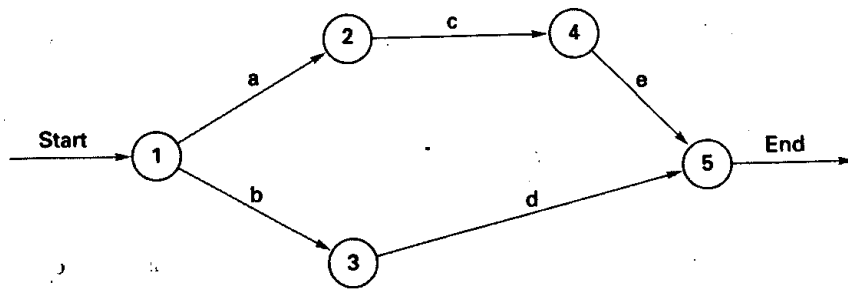
DETERMINISTIC TIME ESTIMATES

The main determinant of the way PERT and CPM networks are analyzed and interpreted is whether activity time estimates are *probabilistic* or *deterministic*. If time estimates can be made with a high degree of confidence that actual times are fairly certain, we say the estimates are **deterministic**. If actual times are subject to variation, we say the estimates are **probabilistic**. Probabilistic time estimates must include an indication of the extent of probable variation.

deterministic Time estimates that are fairly certain.

probabilistic Estimates of times that allow for variation.

This section describes analysis of networks with deterministic time estimates. A later section deals with probabilistic times.

One of the best ways to gain an understanding of the nature of network analysis is to consider a simple example.

EXAMPLE 1

Given the additional information on the bank network of Figure 17.4 shown in Figure 17.5, determine

FIGURE 17.5

AOA diagram
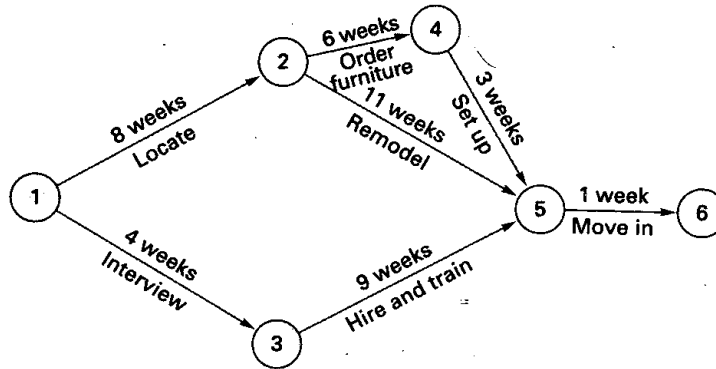
a. The length of each path.
b. The critical path.
c. The expected length of the project.
d. The amount of slack time for each path.

SOLUTION

a. As shown in the following table, the path lengths are 18 weeks, 20 weeks, and 14 weeks.
b. Path 1-2-5-6 is the longest path (20 weeks), so it is the critical path.
c. The expected length of the project is equal to the length of the critical path (i.e., 20 weeks).
d. We find the slack for each path by subtracting its length from the length of the critical path, as shown in the last column of the table. (*Note:* It is sometimes desirable to know the slack time associated with activities. The next section describes a method for obtaining those slack times.)

Path	Length (weeks)	Slack
1-2-4-5-6	$8 + 6 + 3 + 1 = 18$	$20 - 18 = 2$
1-2-5-6	$8 + 11 + 1 = 20*$	$20 - 20 = 0$
1-3-5-6	$4 + 9 + 1 = 14$	$20 - 14 = 6$

*Critical path length.

A COMPUTING ALGORITHM

Many real-life project networks are much larger than the simple network illustrated in the preceding example; they often contain hundreds or even thousands of activities. Because the necessary computations can become exceedingly complex and time-consuming, large networks are generally analyzed by computer programs instead of manually. Planners use an algorithm to develop four pieces of information about the network activities:

ES, the earliest time activity can start, assuming all preceding activities start as early as possible.

EF, the earliest time the activity can finish.

LS, the latest time the activity can start and not delay the project.

LF, the latest time the activity can finish and not delay the project.

Once these values have been determined, they can be used to find

1. Expected project duration.
2. Slack time.
3. The critical path.

Activity-on-Arrow

The three examples that follow illustrate how to compute those values using the precedence diagram of Example 1.

Compute the earliest starting time and earliest finishing time for each activity in the diagram shown in Figure 17.5.

EXAMPLE 2

SOLUTION

Begin by placing brackets at the two ends of each starting activity:

We want to determine and place in the brackets for each activity the earliest starting time, ES, and the earliest finishing time, EF, for every activity, and put them in brackets, as follows:
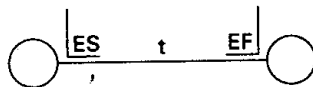
Do this for all activities, beginning at the left side of the precedence diagram and moving to the right side.

Once ES has been determined for each activity, EF can be found by adding the activity time, t, to ES: $ES + t = EF$.

Use an ES of 0 for all starting activities. Thus, activities 1-2 and 1-3 are assigned ES values of 0. This permits computation of the EF for each of these activities:

$$EF_{1\text{-}2} = 0 + 8 = 8 \quad \text{and} \quad EF_{1\text{-}3} = 0 + 4 = 4$$
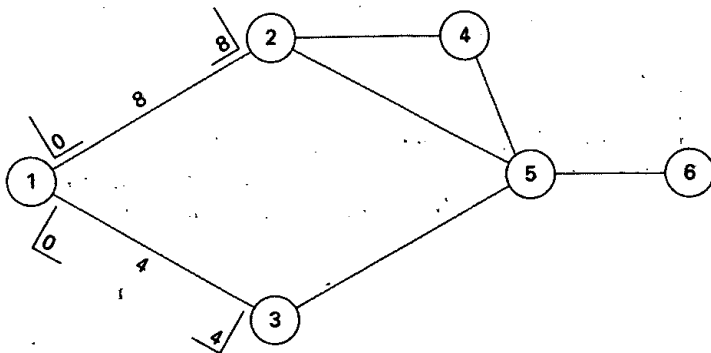
The EF time for an activity becomes the ES time for the next activity to follow it in the diagram. Hence, because activity 1-2 has an EF time of 8, both activities 2-4 and 2-5 have ES times of 8. Similarly, activity 3-5 has an ES time of 4.
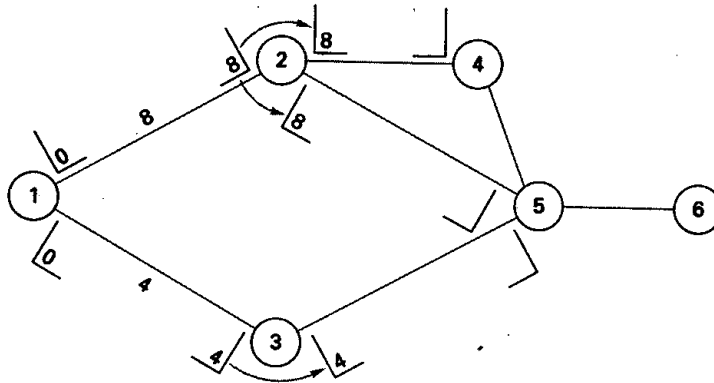
This permits calculation of the EF times for these activities: $EF_{2-4} = 8 + 6 = 14$; $EF_{2-5} = 8 + 11 = 19$; and $EF_{3-5} = 4 + 9 = 13$.
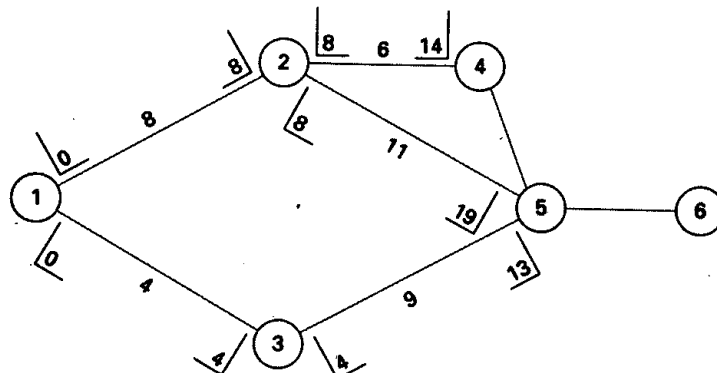
The ES for activity 4-5 is the EF time of activity 2-4, which is 14. Using this value, we find the EF for activity 4-5 is 17; $EF_{4-5} = 14 + 3 = 17$.
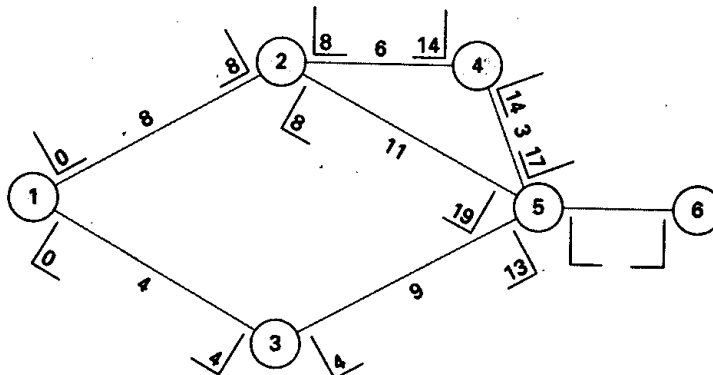
In order to determine the ES for activity 5-6, we must realize that activity 5-6 cannot start until *every* activity that precedes it is finished. Therefore, the *largest* of the EF times for the three activities that precede activity 5-6 determines ES_{5-6}. Hence, the ES for activity 5-6 is 19.
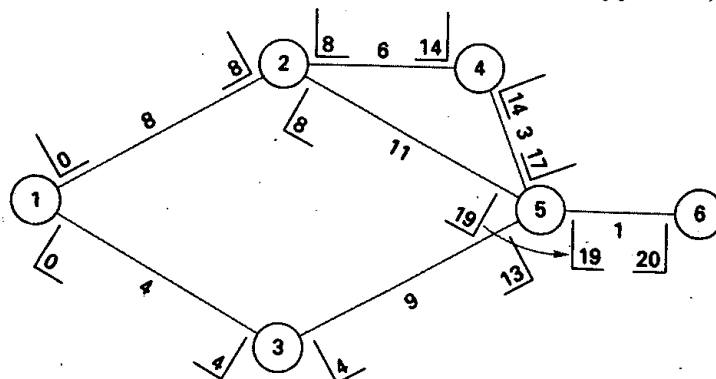
Then the EF for the last activity, 5-6, is 20; $EF_{5-6} = 19 + 1 = 20$. Note that the latest EF is the project duration. Thus, the expected length of the project is 20 weeks.

Computation of earliest starting and finishing times is aided by two simple rules:

1. The earliest finish time for any activity is equal to its earliest start time plus its expected duration, t:

$$EF = ES + t \qquad (17\text{--}1)$$

2. ES for activities at nodes with one entering arrow is equal to EF of the entering arrow. ES for activities leaving nodes with multiple entering arrows is equal to the largest EF of the entering arrow.

Computation of the latest starting and finishing times is aided by the use of two rules:

1. The latest starting time for each activity is equal to its latest finishing time minus its expected duration:

$$LS = LF - t \qquad (17\text{--}2)$$

2. For nodes with one leaving arrow, LF for arrows entering that node equals the LS of the leaving arrow. For nodes with multiple leaving arrows, LF for arrows entering that node equals the smallest LS of leaving arrows.

Finding ES and EF times involves a *forward pass* through the network; finding LS and LF times involves a *backward pass* through the network. Hence, we must begin with the EF of the last activity and use that time as the LF for the last activity. Then we obtain the LS for the last activity by subtracting its expected duration from its LF.

Compute the latest finishing and starting times for the precedence diagram developed in Example 2.

EXAMPLE 3

SOLUTION ————

We must add the LS and LF times to the brackets on the diagram.
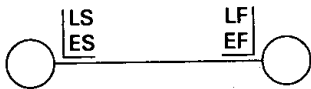
Begin by setting the LF time of the last activity equal to the EF of that activity. Thus,

$$LF_{5-6} = EF_{5-6} = 20 \text{ weeks}$$

Obtain the LS time for activity 5-6 by subtracting the activity time, t, from the LF time:

$$LS_{5-6} = LF_{5-6} - t = 20 - 1 = 19$$

Mark these values on the diagram:
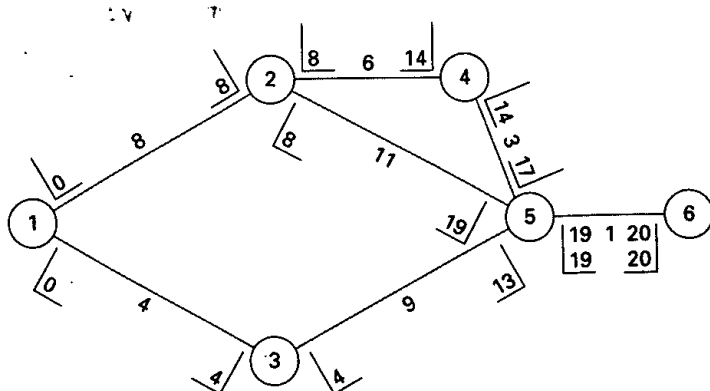
The LS time of 19 for activity 5-6 now becomes the LF time for each of the activities that precedes activity 5-6. This permits determination of the LS times for each of those activities: Subtract the activity time from the LF to obtain the LS time for the activity. The LS time for activity 3-5 is $19 - 9 = 10$.
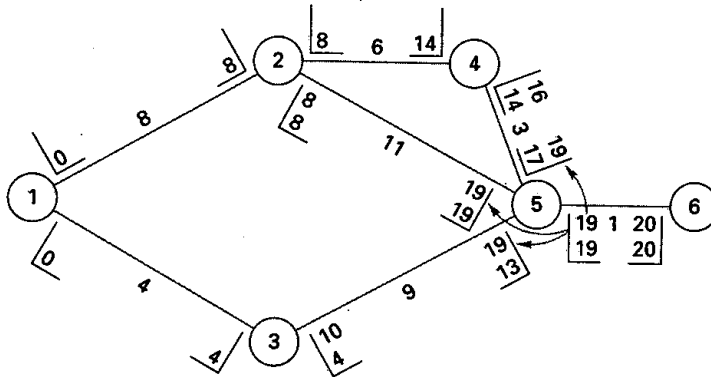
Next, the LS for activity 4-5, which is 16, becomes the LF for activity 2-4, and the LS for activity 3-5, which is 10, becomes the LF for activity 1-3. Using these values, you find the LS for each of these activities by subtracting the activity time from the LF time.
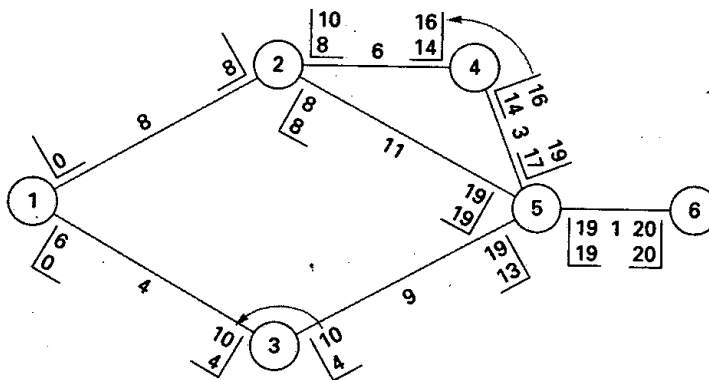
The LF for activity 1-2 is the *smaller* of the two LS times of the activities that 1-2 precedes. Hence, the LF time for activity 1-2 is 8. The reason you use the smaller time is that activity 1-2 must finish at a time that permits all following activities to start no later than their LS times.

Once you have determined the LF time of activity 1-2, find its LS time by subtracting the activity time of 8 from the LF time of 8. Hence, the LS time is 0.
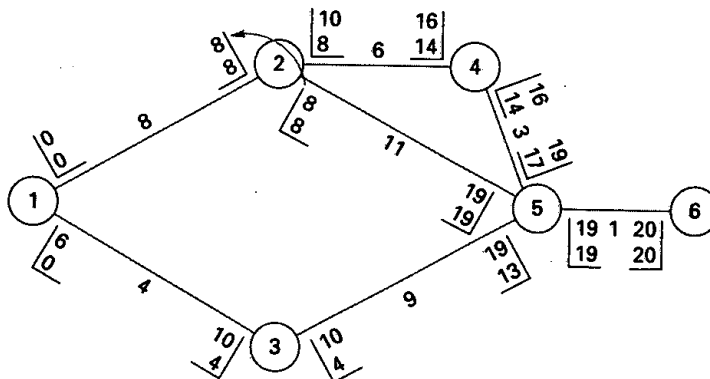
RULES FOR THE COMPUTING ALGORITHM

(Note: For an AON diagram, if a starting node or ending node does not have a time associated with it, ignore that node.)

Forward Pass

For each path, start at the left side of the diagram and work toward the right side.

For each beginning activity: ES = 0.

For each activity: ES + Activity time = EF.

For the following activity: ES = EF of preceding activity.

Note: If an activity has multiple immediate preceding activities, set its ES equal to the largest EF of its immediate predecessors.

Backward Pass

For each path, start at the right side of the diagram and work toward the left side.

Use the largest EF as the LF for all ending activities.

For each activity: LS = LF − Activity time.

For the preceding activity: LF = LS of following activity.

Note: If an activity has multiple immediately following activities, set the activity's LF equal to the smallest LS of the following activities.

Activity-on-Node

The computing algorithm is performed in essentially the same manner in the AON approach. Figure 17.6 shows the node diagram and Figures 17.7 a, b, and c illustrate the computing algorithm.
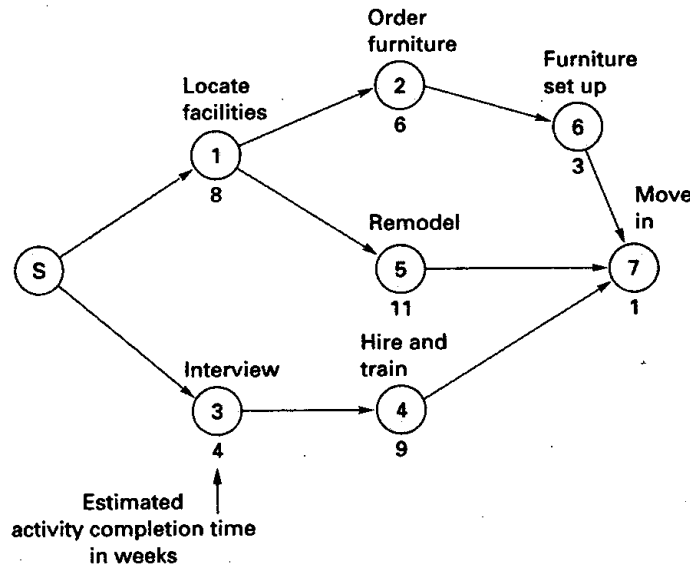
FIGURE 17.6

AON diagram
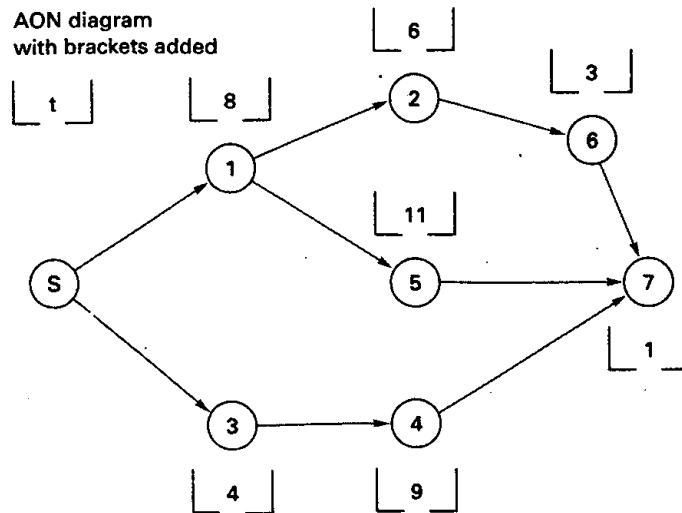
FIGURE 17.7A

FIGURE 17.7B

Forward Pass

$$\boxed{\text{ES} \ t \ \text{EF}}$$

FIGURE 17.7C

Backward Pass

$$\boxed{\begin{matrix} \text{LS} \\ \text{ES} \end{matrix} \ t \ \begin{matrix} \text{LF} \\ \text{EF} \end{matrix}}$$
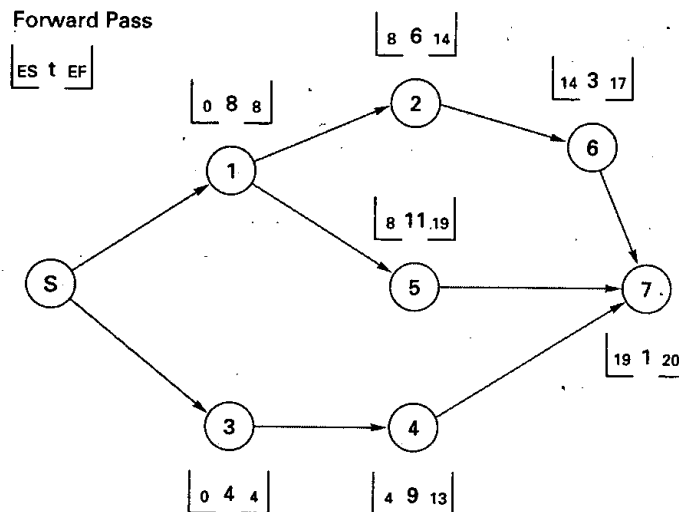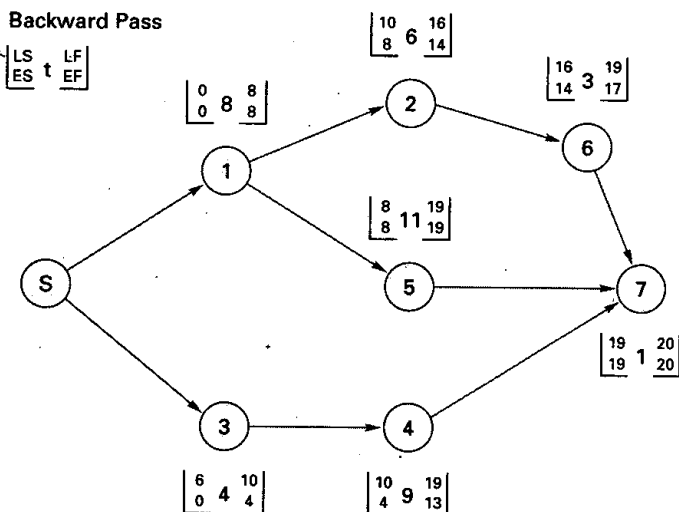
Computing Slack Times

The slack time can be computed in either of two ways:

$$\text{Slack} = \text{LS} - \text{ES} \qquad \text{or} \qquad \text{LF} - \text{EF} \tag{17–3}$$

The critical path using this computing algorithm is denoted by activities with zero slack time. Thus, the table in Example 4 indicates that activities 1-2, 2-5, and 5-6 are all critical activities, which agrees with the results of the intuitive approach demonstrated in Example 1.

Knowledge of slack times provides managers with information for planning allocation of scarce resources and for directing control efforts toward those activities that might be most susceptible to delaying the project. In this regard, it is important to recognize that the activity slack times are based on the assumption that all of the activities on the same path will be started as early as possible and not exceed their expected times. Furthermore, if two activities are both on the same path (e.g., activities 2-4 and 4-5 in the preceding example) and have the same slack (e.g., two weeks), this will be the *total* slack available to both. In essence, the activities have *shared slack*. Hence, if the first activity uses all the slack, there will be zero slack for all following activities on that same path.

TABLE 17.3
Computer printout

| | | SCHEDULE | | | | |
| | | EARLY | | LATE | | |
Activity	Time	ES	EF	LS	LF	Slack
1-2	8.00	0.00	8.00	0.00	8.00	0.00
1-3	4.00	0.00	4.00	6.00	10.00	6.00
2-4	6.00	8.00	14.00	10.00	16.00	2.00
2-5	11.00	8.00	19.00	8.00	19.00	0.00
3-5	9.00	4.00	13.00	10.00	19.00	6.00
4-5	3.00	14.00	17.00	16.00	19.00	2.00
5-6	1.00	19.00	20.00	19.00	20.00	0.00

THE CRITICAL PATH SEQUENCE IS:

SNODE	FNODE	TIME
1	2	8.00
2	5	11.00
5	6	1.00
		20.00

Compute slack times for the preceding example.

EXAMPLE 4

SOLUTION

Either the starting times or the finishing times can be used. Suppose we choose the starting times. Using ES times computed in Example 2 and LS times computed in Example 3, slack times are

Activity	LS	ES	(LS – ES) Slack
1-2	0	0	0
1-3	6	0	6
2-4	10	8	2
2-5	8	8	0
3-5	10	4	6
4-5	16	14	2
5-6	19	19	0

Activities that have a slack of zero are on the critical path. Hence, the critical path is 1-2-5-6.

As noted earlier, this algorithm lends itself to computerization. A computer printout for this problem would appear something like the one shown in Table 17.3.

PROBABILISTIC TIME ESTIMATES

The preceding discussion assumed that activity times were known and not subject to variation. While that condition exists in some situations, there are many others where it does not. Consequently, those situations require a probabilistic approach.

The probabilistic approach involves *three* time estimates for each activity instead of one:

1. **Optimistic time**: The length of time required under optimum conditions; represented by t_o.

2. **Pessimistic time**: The length of time required under the worst conditions; represented by t_p.

3. **Most likely time**: The most probable amount of time required; represented by t_m.

optimistic time The length of time required under optimal conditions.

pessimistic time The length of time required under the worst conditions.

most likely time The most probable length of time that will be required.

A beta distribution is used to describe probabilistic time estimates
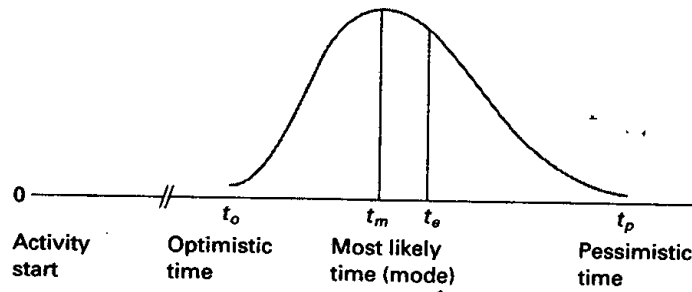
beta distribution Used to describe the inherent variability in activity time estimates.

Managers or others with knowledge about the project can make these time estimates.

The **beta distribution** is generally used to describe the inherent variability in time estimates (see Figure 17.8). Although there is no real theoretical justification for using the beta distribution, it has certain features that make it attractive in practice: The distribution can be symmetrical or skewed to either the right or the left according to the nature of a particular activity; the mean and variance of the distribution can be readily obtained from the three time estimates listed above; and the distribution is unimodal with a high concentration of probability surrounding the most likely time estimate.

Of special interest in network analysis are the average or expected time for each activity, t_e, and the variance of each activity time, σ_i^2. The expected time of an activity, t_e, is a weighted average of the three time estimates:

$$t_e = \frac{t_o + 4t_m + t_p}{6} \tag{17-4}$$

The expected duration of a path (i.e., the path mean) is equal to the sum of the expected times of the activities on that path:

Path mean = Σ of expected times of activities on the path. $\qquad$ (17-5)

The standard deviation of each activity's time is estimated as one-sixth of the difference between the pessimistic and optimistic time estimates. (Analogously, essentially all of the area under a normal distribution lies within three standard deviations of the mean, which is a range of six standard deviations.) We find the variance by squaring the standard deviation. Thus,

$$\sigma^2 = \left[\frac{(t_p - t_o)}{6}\right]^2 \quad \text{or} \quad \frac{(t_p - t_o)^2}{36} \tag{17-6}$$

The size of the variance reflects the degree of uncertainty associated with an activity's time: The larger the variance, the greater the uncertainty.

It is also desirable to compute the standard deviation of the expected time for *each path*. We can do this by summing the variances of the activities on a path and then taking the square root of that number; that is,

$$\sigma_{path} = \sqrt{\Sigma(\text{variances of activities on path})} \tag{17-7}$$

Example 5 illustrates these computations.

EXAMPLE 5

The network diagram for a project is shown in the accompanying figure, with three time estimates for each activity. Activity times are in weeks. Do the following:

a. Compute the expected time for each activity and the expected duration for each path.

b. Identify the critical path.

c. Compute the variance of each activity and the variance and standard deviation of each path.

AOA diagram
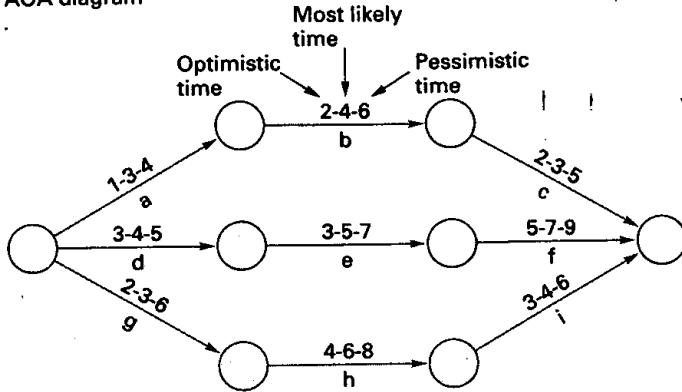
AON diagram
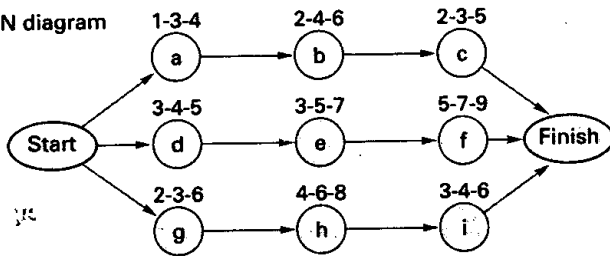
SOLUTION

a.

Path	Activity	t_o	t_m	t_p	$t_e = \dfrac{t_o + 4t_m + t_p}{6}$	Path Total
	a	1	3	4	2.83	
a-b-c	b	2	4	6	4.00	10.00
	c	2	3	5	3.17	
	d	3	4	5	4.00	
d-e-f	e	3	5	7	5.00	16.00
	f	5	7	9	7.00	
	g	2	3	6	3.33	
g-h-i	h	4	6	8	6.00	13.50
	i	3	4	6	4.17	

Header above t_o, t_m, t_p columns: TIMES

b. The path that has the longest expected duration is the critical path. Because path d-e-f has the largest path total, it is the critical path.

c.

Path	Activity	t_o	t_m	t_p	$\sigma^2_{act} = \dfrac{(t_p - t_o)^2}{36}$	σ^2_{path}	σ_{path}
a-b-c	a	1	3	4	$(4-1)^2/36 = 9/36$		
	b	2	4	6	$(6-2)^2/36 = 16/36$	$34/36 = 0.944$	0.97
	c	2	3	5	$(5-2)^2/36 = 9/36$		
d-e-f	d	3	4	5	$(5-3)^2/36 = 4/36$		
	e	3	5	7	$(7-3)^2/36 = 16/36$	$36/36 = 1.00$	1.00
	f	5	7	9	$(9-5)^2/36 = 16/36$		
g-h-i	g	2	3	6	$(6-2)^2/36 = 16/36$		
	h	4	6	8	$(8-4)^2/36 = 16/36$	$41/36 = 1.139$	1.07
	i	3	4	6	$(6-3)^2/36 = 9/36$		

Header above t_o, t_m, t_p columns: TIMES

FIGURE 17.9

Activity distributions and the path distribution
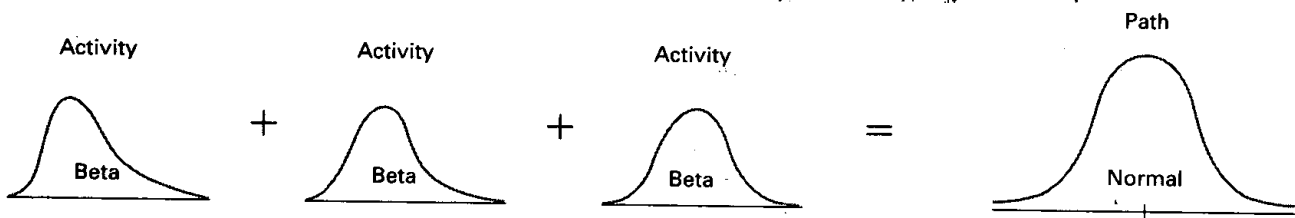
Knowledge of the expected path times and their standard deviations enables a manager to compute probabilistic estimates of the project completion time, such as these:

The probability that the project will be completed by a specified time.

The probability that the project will take longer than its scheduled completion time.

These estimates can be derived from the probability that various paths will be completed by the specified time. This involves the use of the normal distribution. Although activity times are represented by a beta distribution, the path distribution is represented by a normal distribution. The summing of activity times (random variables) results in a normal distribution. This is illustrated in Figure 17.9. The rationale for using a normal distribution is that sums of random variables (activity times) will tend to be normally distributed, regardless of the distributions of the variables. The normal tendency improves as the number of random variables increases. However, even when the number of items being summed is fairly small, the normal approximation provides a reasonable approximation to the actual distribution.

DETERMINING PATH PROBABILITIES

The probability that a given path will be completed in a specified length of time can be determined using the following formula:

$$z = \frac{\text{Specified time} - \text{Path mean}}{\text{Path standard deviation}}$$

(17–8)

The resulting value of z indicates how many standard deviations of the path distribution the specified time is beyond the expected path duration. The more positive the value, the better. (A negative value of z indicates that the specified time is *earlier* than the expected path duration.) Once the value of z has been determined, it can be used to obtain the probability that the path will be completed by the specified time from Appendix B, Table B. Note that the probability is equal to the area under the normal curve to the left of z, as illustrated in Figure 17.10.

FIGURE 17.10

The path probability is the area under a normal curve to the left of z
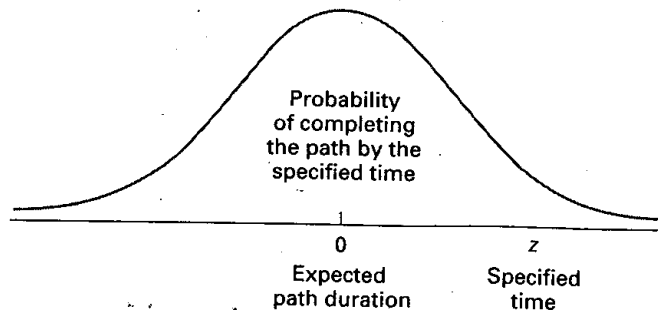
If the value of z is $+3.00$ or more, the path probability is close to 100 percent (for $z = +3.00$, it is .9987). Hence, it is very likely the activities that make up the path will be completed by the specified time. For that reason, a useful rule of thumb is to treat the path probability as being equal to 100 percent if the value of z is $+3.00$ or more.

> Rule of thumb: If the value of z is $+3.00$ or more, treat the probability of path completion by the specified time as 100 percent.

A project is not completed until *all* of its activities have been completed, not only those on the critical path. It sometimes happens that another path ends up taking more time to complete than the critical path, in which case the project runs longer than expected. Hence, it can be risky to focus exclusively on the critical path. Instead, one must consider the possibility that at least one other path will delay timely project completion. This requires determining the probability that *all* paths will finish by a specified time. To do that, find the probability that each path will finish by the specified time, and then multiply those probabilities. The result is the probability that the *project* will be completed by the specified time.

It is important to note the assumption of **independence**. It is assumed that path duration times are independent of each other. In essence, this requires two things: Activity times are independent of each other, and each activity is only on one path. For activity times to be independent, the time for one must not be a function of the time of another; if two activities were always early or late together, they would not be considered independent. The assumption of independent *paths* is usually considered to be met if only a *few* activities in a large project are on multiple paths. Even then, common sense should govern the decision of whether the independence assumption is justified.

independence Assumption that path duration times are independent of each other; requiring that activity times be independent, and that each activity is on only one path.

Using the information from Example 5, answer the following questions:

EXAMPLE 6

a. Can the paths be considered independent? Why?

b. What is the probability that the project can be completed within 17 weeks of its start?

c. What is the probability that the project will be completed within 15 weeks of its start?

d. What is the probability that the project will *not* be completed within 15 weeks of its start?

SOLUTION

a. Yes, the paths can be considered independent, since no activity is on more than one path and you have no information suggesting that any activity times are interrelated.

b. To answer questions of this nature, you must take into account the degree to which the path distributions "overlap" the specified completion time. This overlap concept is illustrated in the accompanying figure, which shows the three path distributions, each centered on that path's expected duration, and the specified completion time of 17 weeks. The shaded portion of each distribution corresponds to the probability that the part will be completed within the specified time. Observe that paths a-b-c and g-h-i are well enough to the left of the specified time, so that it is highly likely that both will be finished by week 17, but the critical path overlaps the specified completion time. In such cases, you need consider only the distribution of path d-e-f in assessing the probability of completion by week 17.

To find the probability for a path you must first compute the value of z using Formula 17–8 for the path. For example, for path d-e-f, we have

$$z = \frac{17 - 16}{1.00} = +1.00$$

Turning to Appendix B, Table B with $z = +1.00$, you will find that the area under the curve to the left of z is .8413. The computations are summarized in the following table. *Note:* If the value of z exceeds $+3.00$, treat the probability of completion as being equal to 1.000.

Path	$z = \dfrac{17 - \text{Expected path duration}}{\text{Path standard deviation}}$	Probability of Completion in 17 Weeks
a-b-c	$\dfrac{17 - 10}{0.97} = +7.22$	1.00
d-e-f	$\dfrac{17 - 16}{1.00} = +1.00$	.8413
g-h-i	$\dfrac{17 - 13.5}{1.07} = +3.27$	1.00

$P(\text{finish by week 17}) = P(\text{path a-b-c finish}) \times P(\text{path d-e-f finish}) \times P(\text{path g-h-i finish})$

$$= \quad 1.00 \quad \times \quad .8413 \quad \times 1.00 = .8413$$

c. For a specified time of 15 weeks, the z values are

Path	$z = \dfrac{15 - \text{Expected path duration}}{\text{Path standard deviation}}$	Probability of Completion in 15 Weeks
a-b-c	$\dfrac{15 - 10.00}{0.97} = +5.15$	1.00
d-e-f	$\dfrac{15 - 16.00}{1.00} = -1.00$	.1587
g-h-i	$\dfrac{15 - 13.50}{1.07} = +1.40$	.9192

Paths d-e-f and g-h-i have z values that are less than $+3.00$.

From Appendix B, Table B, the area to the *left* of $z = -1.00$ is .1587, and the area to the *left* of $z = +1.40$ is .9192. The path distributions are illustrated in the figure. The joint probability of all finishing before week 15 is the product of their probabilities: $1.00(.1587)(.9192) = .1459$.

d. The probability of not finishing before week 15 is the complement of the probability obtained in part *c*: $1 - .1459 = .8541$.
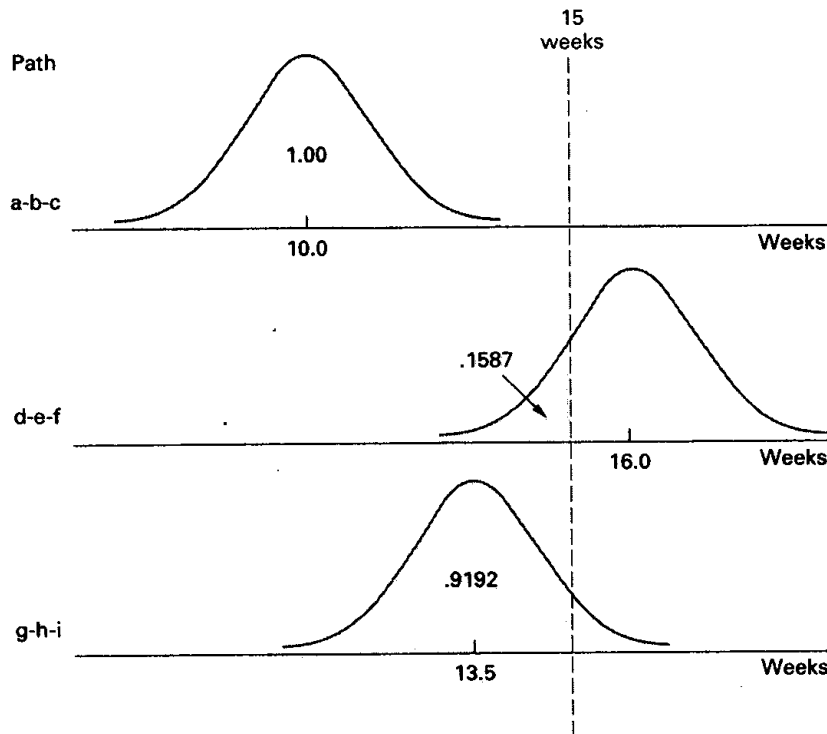
SIMULATION

We have examined a method for computing the probability that a project would be completed in a specified length of time. That discussion assumed that the paths of the project were *independent*; that is, the same activities are not on more than one path. If an activity were on more than one path and it happened that the completion time for that activity far exceeded its expected time, all paths that included that activity would be affected and, hence, their times would not be independent. Where activities are on multiple paths, one must consider if the preceding approach can be used. For instance, if only a few activities are on multiple paths, particularly if the paths are *much* shorter than the critical path, that approach may still be reasonable. Moreover, for purposes of illustration, as in the text problems and examples, the paths are treated as being independent when, in fact, they may not be.

In practice, when *dependent* cases occur, project planners often use *simulation*. It amounts to a form of repeated sampling wherein many passes are made through the project network. In each pass, a randomly selected value for each activity time is made based on the characteristics of the activity's probability distribution (e.g., its mean, standard deviation, and distribution type). After each pass, the expected project duration is determined by adding the times along each path and designating the time of the longest path as the project duration. After a large number of such passes (e.g., several hundred), there is enough information to prepare a frequency distribution of the project duration times. Planners can use this distribution to make a probabilistic assessment of the actual project duration, allowing for some activities that are on more than one path. Problem 19 in the supplement to Chapter 18 illustrates this.

TIME–COST TRADE-OFFS: CRASHING

Estimates of activity times for projects usually are made for some given level of resources. In many situations, it is possible to reduce the length of a project by injecting additional resources. The impetus to shorten projects may reflect efforts to avoid late penalties, to take advantage of monetary incentives for timely or early competition of a project, or to free

Projects such as bridge construction involve large budgets, many suppliers and employees, and backup plans. Materials, people, processes, and uncontrollable circumstances, such as weather, can cause delays and additional costs.
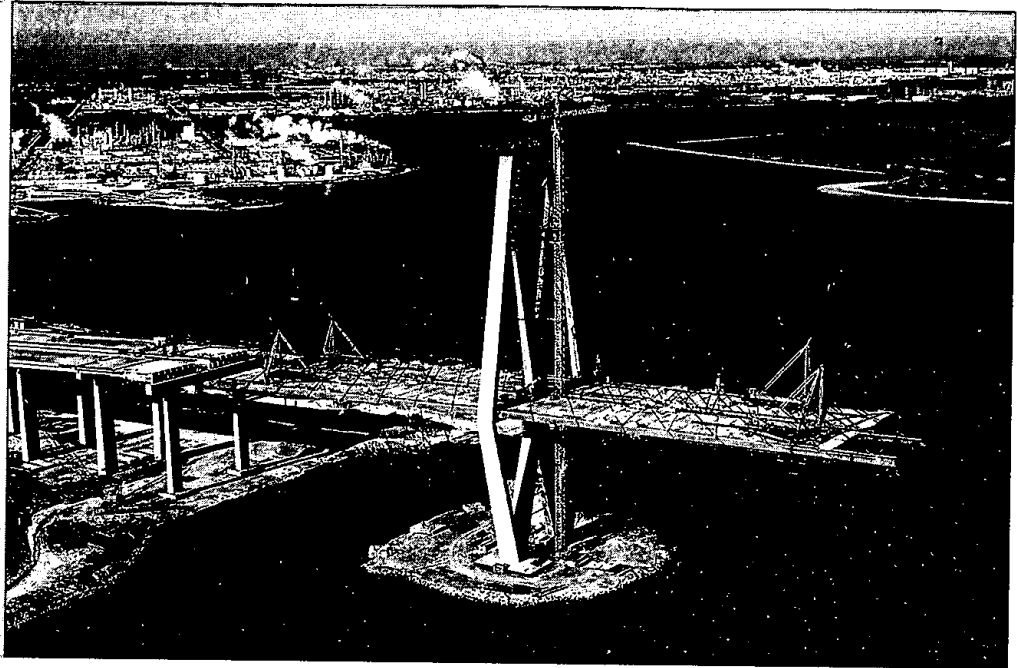
crash Shortening activity durations.

resources for use on other projects. In new product development, shortening may lead to a strategic benefit: beating the competition to the market. In some cases, however, the desire to shorten the length of a project merely reflects an attempt to reduce the indirect costs associated with running the project, such as facilities and equipment costs, supervision, and labor and personnel costs. Managers often have various options at their disposal that will allow them to shorten, or **crash**, certain activities. Among the most obvious options are the use of additional funds to support additional personnel or more efficient equipment, and the relaxing of some work specifications. Hence, a project manager may be able to shorten a project by increasing *direct* expenses to speed up the project, thereby realizing savings on indirect project costs. The goal in evaluating time–cost trade-offs is to identify activities that will reduce the sum of the indirect and direct project costs.

In order to make a rational decision on which activities, if any, to crash and on the extent of crashing desirable, a manager needs certain information:

1. Regular time and crash time estimates for each activity.
2. Regular cost and crash cost estimates for each activity.
3. A list of activities that are on the critical path.

Activities on the critical path are potential candidates for crashing, because shortening noncritical activities would not have an impact on total project duration. From an economic standpoint, activities should be crashed according to crashing costs: Crash those with the lowest crash costs first. Moreover, crashing should continue as long as the cost to crash is less than the benefits derived from crashing. Figure 17.11 illustrates the basic cost relationships.

Crashing analysis requires estimates of regular and crash times and costs for each activity, path lengths, and identification of critical activities. The general procedure for crashing is

1. Crash the project one period at a time.
2. Crash the least expensive activity that is on the critical path.
3. When there are multiple critical paths, find the sum of crashing the least expensive activity on each critical path. If two or more critical paths share common activities, compare the least expensive cost of crashing a common activity shared by critical paths with the sum for the separate critical paths.
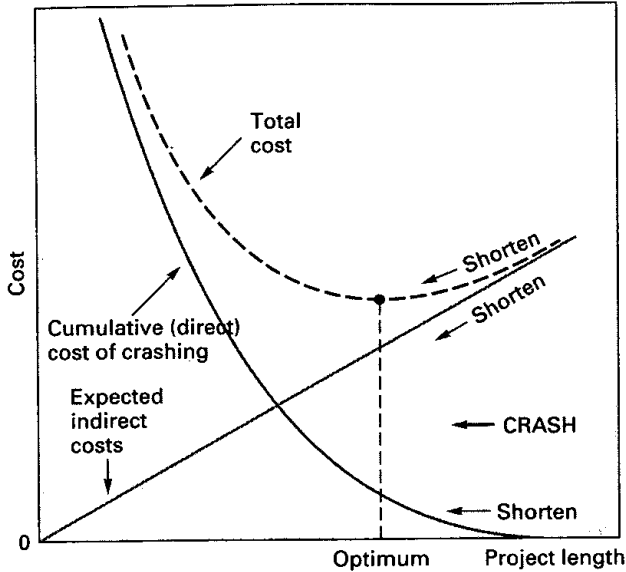
FIGURE 17.11

Crashing activities*

* Crashing activities reduce indirect project costs and
increase direct costs; the optimum amount of crashing
results in minimizing the sum of these two types of costs. .

Using the following information, develop the optimal time–cost solution. Indirect project costs
are $1,000 per day.

EXAMPLE 7

Activity	Normal Time	Crash Time	Cost per Day to Crash
a	6	6	—
b	10	8	$500
c	5	4	300
d	4	1	700
e	9	7	600
f	2	1	800
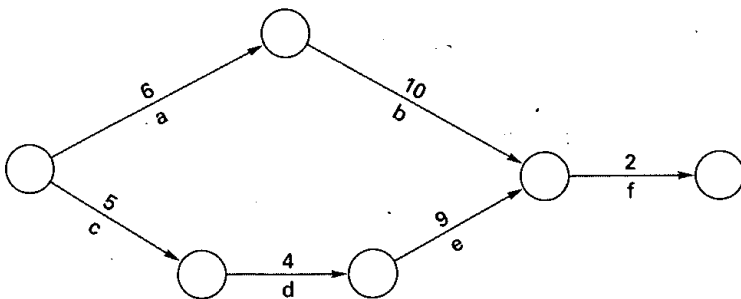
SOLUTION ——————

a.   Determine which activities are on the critical path, its length, and the length of the other
     path: .

Path	Length
a-b-f	18
c-d-e-f.....	20 (critical path)

b.   Rank the critical path activities in order of lowest crashing cost, and determine the num-
     ber of days each can be crashed.

Part Seven   Project Management

Activity	Cost per Day to Crash	Available Days
c...........	$300	1
e...........	600	2
d...........	700	3
f...........	800	1

c. Begin shortening the project, one day at a time, and check after each reduction to see which path is critical. (After a certain point, another path may equal the length of the shortened critical path.) Thus:

(1) Shorten activity c one day at a cost of $300. The length of the critical path now becomes 19 days.

(2) Activity c cannot be shortened any more. Shorten activity e one day at a cost of $600. The length of path c-d-e-f now becomes 18 days, which is the same as the length of path a-b-f.

(3) The paths are now both critical; further improvements will necessitate shortening both paths.

The remaining activities for crashing and their costs are:

Path	Activity	Crash Cost (per day)
a-b-f	a	No reduction possible
	b	$500
	f	800
c-d-e-f	c	No further reduction possible
	d	$700
	e	600
	f	800

At first glance, it would seem that crashing f would not be advantageous, because it has the highest crashing cost. However, f is on *both* paths, so shortening f by one day would shorten *both* paths (and hence, the project) by one day for a cost of $800. The option of shortening the least expensive activity on each path would cost $500 for b and $600 for e, or $1,100. Thus shorten f by one day. The project duration is now 17 days.

(4) At this point, no additional improvement is feasible. The cost to crash b is $500 and the cost to crash e is $600, for a total of $1,100, and that would exceed the indirect project costs of $1,000 per day.

(5) The crashing sequence is summarized below:

	LENGTH AFTER CRASHING n DAYS:			
Path	$n = 0$	1	2	3
a-b-f	18	18	18	17
c-d-e-f	20	19	18	17
Activity crashed		c	e	f
Cost		$300	$600	$800

An important benefit of the sequential crashing procedure just described is that it provides the ability to quote different budget costs for different project times.

ADVANTAGES OF USING PERT AND POTENTIAL SOURCES OF ERROR

PERT and similar project scheduling techniques can provide important services for the project manager. Among the most useful features are these:

1. Use of these techniques forces the manager to organize and quantify available information and to recognize where additional information is needed.

2. The techniques provide a graphic display of the project and its major activities.

3. They identify (a) activities that should be closely watched because of the potential for delaying the project and (b) other activities that have slack time and so can be delayed without affecting project completion time. This raises the possibility of reallocating resources to shorten the project.

No analytical technique is without potential errors. Among the more important sources of errors are

1. When developing the project network, managers may unwittingly omit one or more important activities.

2. Precedence relationships may not all be correct as shown.

3. Time estimates may include a fudge factor; managers may feel uncomfortable about making time estimates because they appear to commit themselves to completion within a certain time period.

4. There may be a tendency to focus solely on activities that are on the critical path. As the project progresses, other paths may become critical. Furthermore, major risk events may not be on the critical path.

TECHNOLOGY FOR MANAGING PROJECTS

Technology has had a number of benefits for project management. Among those benefits are the use of computer-aided design (CAD) to produce updated prototypes on construction and product-development projects; software such as Lotus Notes to keep team members who are in separate locations in close contact; the ability for remote viewing of projects, allowing those in different locations a firsthand view of progress and problems; and computer software.

Project management software is often used to facilitate project management. Here are a few of the many project management software packages that are available: CA-Super Project, Harvard Total Project Manager, Microsoft Project for Windows, and Time Line. The particular version, e.g., 4.0, is not included because new versions are released frequently.

There are many advantages to using a project management software package. Among them are the following:

- It imposes a methodology and a common project management terminology.
- It provides a logical planning structure.
- It can enhance communication among team members.
- It can flag the occurrence of constraint violations.
- It automatically formats reports.
- It can generate multiple levels of summary reports and detailed reports.
- It enables "what-if" scenarios.
- It can generate various chart types, including basic Gantt charts.

Certain requirements are associated with using project management software. One is the need to keep the system current by entering changes as soon as possible. Also, as with any software, as new versions are released, the vendor may drop support for older versions, requiring users to move to the latest version, and that can mean learning to use that version.

One thing to keep in mind is that project management is more than choosing the right software. There is much that a project manager must do. Recall the key decisions that were discussed early in the chapter.

Operations Strategy

Projects can present both strategic opportunities and strategic risks, so it is critical for management to devote adequate attention and resources to projects.

Projects are often used in situations that have some degree of uncertainty, which can result in delays, budget overruns, and failure to meet technical requirements. To minimize the impact of these possibilities, management must ensure that careful planning, wise selection of project managers and team members, and monitoring of the project occur.

Computer software and tools such as PERT can greatly assist project management. However, care must be taken to avoid focusing exclusively on the critical path. The obvious reason is that as the project progresses, other paths may become critical. But another, less obvious, reason is that key risk events may not be on the critical path. Even so, if they occur, they can have a major impact on the project.

Some projects can benefit from *managing slack.* Activities are monitored according to the percentage complete, so that if activities finish early, following activities can be started early. This may permit the project to finish early, or at least conserve slack to offset problems that might arise later in the project.

Another approach to managing slack is suggested by Goldratt in his book, *Critical Chain.*[2] He asserts that people have a tendency to build in extra time in estimated times, and, by insisting that estimates be more realistic, managers can get a more realistic idea of the overall estimated time for a project. Goldratt recognizes that time buffers can be useful in certain instances, so he recommends carefully analyzing the project to identify problem areas, and inserting time buffers for those activities.

RISK MANAGEMENT

Risks are inherent in projects. They relate to the occurrence of events that can have undesirable consequences, such as delays, increased costs, and an inability to meet technical specifications. In some instances, there is the risk that events will occur that will cause a project to be terminated. Although careful planning can reduce risks, no amount of planning can eliminate chance events due to unforeseen, or uncontrollable, circumstances.

The probability of occurrence of risk events is highest near the beginning of a project and the lowest near the end. However, the cost associated with risk events tends to be lowest near the beginning of a project and highest near the end. (See Figure 17.12.)

Good risk management entails identifying as many potential risks as possible, analyzing and assessing those risks, working to minimize the probability of their occurrence, and establishing contingency plans (and funds) for dealing with any that do occur. Much of this takes place before the start of a project, although it is not unusual for this process to be repeated during the project as experience grows and new information becomes available.

The first step is to identify the risks. Typically, there are numerous sources of risks, although the more experience an organization has with a particular type of work, the fewer and more identifiable the risks. Everyone associated with the project should have responsibility for the task of identifying risks. Brainstorming sessions and questionnaires can be useful in this regard.

Once risks have been identified, each risk must be evaluated to determine its probability of occurrence and the potential consequences if it does occur. Both quantitative and qualitative approaches have merit. Managers and workers can contribute to this effort, and experts might be called on. Experience with previous projects can be useful. Many tools might be applied, including scenario analysis, simulation, and PERT (described earlier in the chapter.)

Risk reduction can take a number of forms. Much depends on the nature and scope of a project. "Redundant" (backup) systems can sometimes be used to reduce the risk of failure. For example, an emergency generator could supply power in the event of an electrical failure. Another approach is frequent monitoring of critical project dimensions with the goal of catching and eliminating problems in their early stages, before they cause extensive damage. Risks can sometimes be transferred, say by outsourcing a particular component of a project. Risk-sharing is an-

[2]Eliyahu M. Goldratt, *Critical Chain* (Great Barrington, MA: North River Press, 1997).
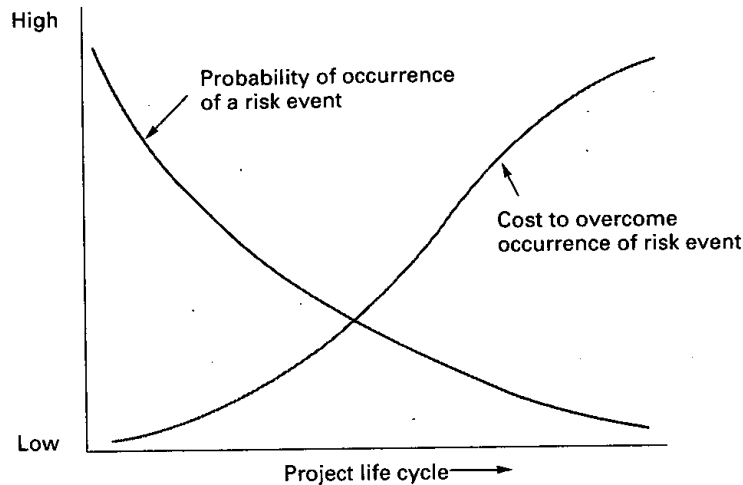
High

Probability of occurrence
of a risk event

Cost to overcome
occurrence of risk event

Low

Project life cycle ⟶

other possibility. This might involve partnering, which can spread risks among partners; this approach may also reduce risk by enlarging the sphere of sources of ideas for reducing the risk.

SUMMARY

Projects are composed of a unique set of activities established to realize a given set of objectives in a limited time span. Projects go through a life cycle that involves definition, planning, execution, and delivery/termination. The nonroutine nature of project activities places a set of demands on the project manager that are different in many respects from those the manager of more routine operations activities experiences, both in planning and coordinating the work and in the human problems encountered. Ethical conduct and risk management are among the key issues project managers must deal with.

PERT and CPM are two commonly used techniques for developing and monitoring projects. Although each technique was developed independently and for expressly different purposes, time and practice have erased most of the original differences, so that now there is little distinction between the two. Either provides the manager with a rational approach to project planning and a graphical display of project activities. Both depict the sequential relationships that exist among activities and reveal to managers which activities must be completed on time to achieve timely project completion. Managers can use that information to direct their attention toward the most critical activities.

Two slightly different conventions can be used for constructing a network diagram. One designates the arrows as activities; the other designates the nodes as activities.

The task of developing and updating project networks quickly becomes complex for projects of even moderate size, so computer software is important. Among the advantages of using project management software are the provision for a logical planning structure, enhanced communication, and automatically formatted charts and reports.

In some instances, it may be possible to shorten, or crash, the length of a project by shortening one or more of the project activities. Typically, such gains are achieved by the use of additional resources, although in some cases, it may be possible to transfer resources among project activities. Generally, projects are shortened to the point where the cost of additional reduction would exceed the benefit of additional reduction, or to a specified time.

KEY TERMS

activities, 737
activity-on-arrow (AOA), 737
activity-on-node (AON), 737
beta distribution, 748
CPM, 736
crash, 754
critical activities, 738
critical path, 738
deterministic, 739
events, 737
independence, 751
most likely time, 747

network (precedence) diagram, 737
optimistic time, 747
path, 738
PERT, 736
pessimistic time, 747
precedence diagram, 737
probabilistic, 739
project champion, 732
projects, 728
slack, 738
work breakdown structure (WBS), 735

Problem 1

The following table contains information related to the major activities of a research project. Use the information to do the following:

a. Draw a precedence diagram using AOA.

b. Find the critical path.

c. Determine the expected length of the project.

Activity	Immediate Predecessor	Expected Time (days)
a............	—	5
c...........	a	8
d...........	c	2
b...........	a	7
e...........	—	3
f...........	e	6
i...........	b, d	10
m..........	f, i	8
g...........	—	1
h..........	g	2
k..........	h	17
end........	k, m	

Solution

a. In constructing networks, these observations can be useful.

(1) Activities with no predecessors are at the beginning (left side) of the network.

(2) Activities with multiple predecessors are located at path intersections.

Complete the diagram in sections. Go down the activity list in order to avoid overlooking any activities.

AOA diagram

AON diagram
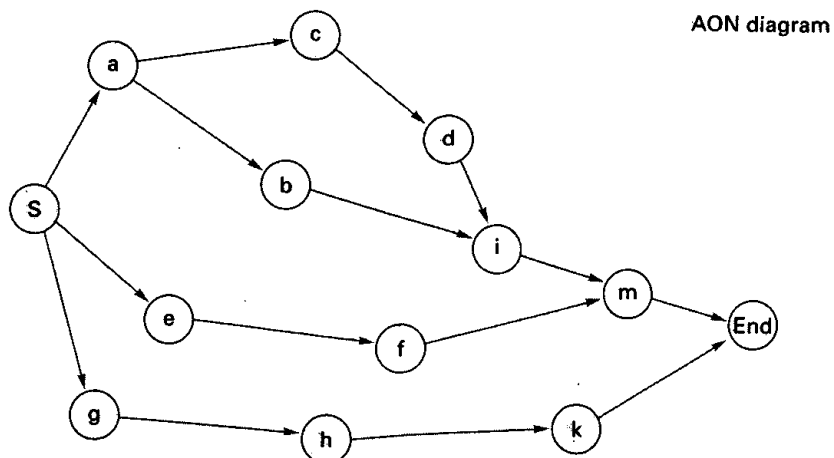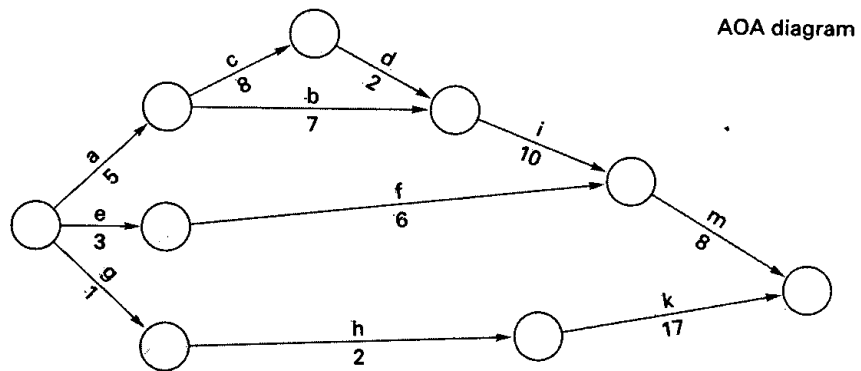
Here are some additional hints for constructing a precedence diagram.

(1) Use pencil.

(2) Start and end with a single node.

(3) Avoid having paths cross each other.

(4) Have activities go from left to right.

(5) Use only one arrow between any pair of nodes.

Yes   No
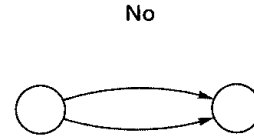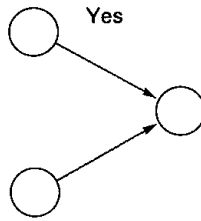
b. and *c.*

Path	Length (days)
a-c-d-i-m*	$5 + 8 + 2 + 10 + 8 = 33^\dagger$
a-b-i-m	$5 + 7 + 10 + 8 = 30$
e-f-m	$3 + 6 + 8 = 17$
g-h-k	$1 + 2 + 17 = 20$

*Critical path.

†Expected project duration.

Using the computing algorithm, determine the slack times for the following AOA diagram. Identify the activities that are on the critical path.
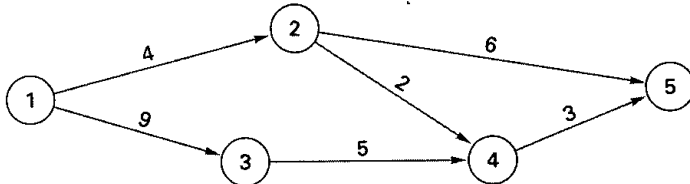
Problem 2

The task of determining ES, EF, LS, and LF times can be greatly simplified by setting up two brackets for each activity, as illustrated:
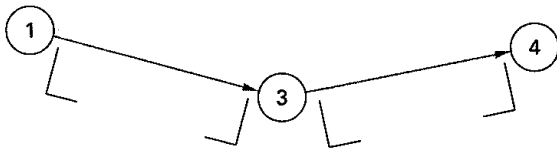
Solution

   The bracket at the left of each activity will eventually be filled in with the earliest and latest *starting* times, and the bracket at the right end of each activity will be filled in with the earliest and latest *finishing* times:

LS   LF

ES   EF

   This is accomplished in a two-step process. First, determine the earliest starting times and earliest finishing times, working from left to right, as shown in the following diagram.
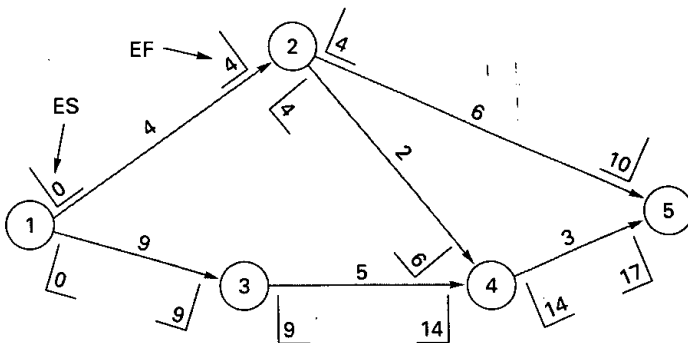
Thus, activity 1-2 can start at 0. With a time of 4, it can finish at $0 + 4 = 4$. This establishes the earliest start for all activities that begin at node 2. Hence, 2-5 and 2-4 can start no earlier than time 4. Activity 2-5 has an early finish of $4 + 6 = 10$, and activity 2-4 has an early finish of $4 + 2 = 6$. At this point, it is impossible to say what the earliest start is for 4-5; that will depend on which activity, 3-4 or 2-4, has the latest EF. Consequently, it is necessary to compute ES and EF along the lower path. Assuming an ES of 0 for activity 1-3, its EF will be 9, so activity 3-4 will have an ES of 9 and an EF of $9 + 5 = 14$.

Considering that the two activities entering node 4 have EF times of 6 and 14, the earliest that activity 4-5 can start is the *larger* of these, which is 14. Hence, activity 4-5 has an ES of 14 and an EF of $14 + 3 = 17$.

Now compare the EFs of the activities entering the final node. The larger of these, 17, is the expected project duration.

The LF and LS times for each activity can now be determined by working backward through the network (from right to left). The LF for the two activities entering node 5 is 17—the project duration. In other words, to finish the project in 17 weeks, these last two activities must both finish by that time.

In the case of activity 4-5, the LS necessary for an LF of 17 is $17 - 3 = 14$. This means that both activities 2-4 and 3-4 must finish no later than 14. Hence, their LF times are 14. Activity 3 has an LS time of $14 - 5 = 9$, making the LF of activity 1-3 equal to 9, and its LS equal to $9 - 9 = 0$.

Activity 2-4, with an LF time of 14, has an LS time of $14 - 2 = 12$. Activity 2-5 has an LF of 17 and therefore an LS of $17 - 6 = 11$. Thus, the latest activity 2-5 can start is 11, and the latest 2-4 can start is 12 in order to finish by week 17. Since activity 1-2 precedes *both* of these activities, it can finish no later than the *smaller* of these, which is 11. Hence, activity 1-2 has an LF of 11 and an LS of $11 - 4 = 7$.

The ES, EF, LF, and LS times are shown on the following network.
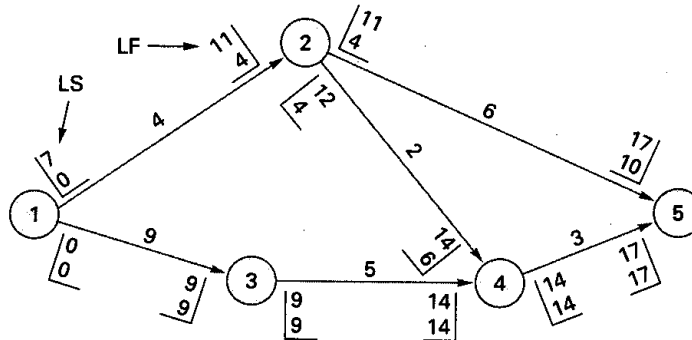
The slack time for any activity is the difference between *either* LF and EF *or* LS and ES. Thus,

Activity	LS	ES	Slack	or	LF	EF	Slack
1-2	7	0	7		11	4	7
2-5	11	4	7		17	10	7
2-4	12	4	8		14	6	8
1-3	0	0	0		9	9	0
3-4	9	9	0		14	14	0
4-5	14	14	0		17	17	0

The activities with zero slack times indicate the critical path. In this case the critical path is 1-3-4-5.

When working problems of this nature, keep in mind the following:

a. The ES time for leaving activities of nodes with multiple entering activities is the largest EF of the entering activities.

b. The LF for an entering activity for nodes with multiple leaving activities is the smallest LS of the leaving activities.

Problem 3

Expected times and variances for the major activities of an R&D project are depicted in the following PERT chart. Determine the probability that project completion time will be

a. 50 weeks or less.

b. More than 50 weeks.
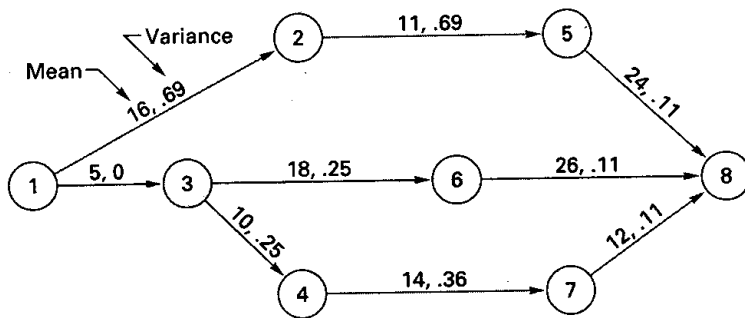
Solution

a. Compute the mean and standard deviation for each path:

Path	Expected Time (weeks)	Standard Deviation (weeks)
1-2-5-8	16 + 11 + 24 = 51	$\sqrt{.69 + .69 + .11} = 1.22$
1-3-6-8	5 + 18 + 26 = 49	$\sqrt{.00 + .25 + .11} = 0.60$
1-3-4-7-8	5 + 10 + 14 + 12 = 41	$\sqrt{.00 + .25 + .36 + .11} = 0.85$

b. Compute the z for each path for the length specified. For any path that has a z of + 3.00 or more, treat its probability of completion before the specified time as 1.00. Use

$$z = \frac{50 - t_{path}}{\sigma_{path}}$$

   The probability that each path will be completed in 50 weeks or less is shown in the corresponding diagram. (Probabilities are from Appendix B, Table B.) The probability that the project will be completed in 50 weeks or less depends on all three paths being completed in that time. Because z for path 1-3-4-7-8 is greater than +3.00, it is treated as having a probability of completion in 50 weeks of 100 percent. It is less certain that the other two paths will be completed in that time. The probability that *both* will not exceed 50 is the *product* of their individual probabilities of completion. Thus, .2061(.9525) (1.00) = .1963.

   The probability that the project *will* exceed 50 weeks is the complement of this number, which is 1.000 − .1963 = .8037. (Note that it is *not* the product of the path probabilities.)

$T = 50$ weeks

$$z = \frac{50 - 51}{1.22} = -.82$$

.2061

-.82   0   z scale

50   51   Weeks

$$z = \frac{50 - 49}{.60} = 1.67$$

.9525

0   1.67   z scale

49   50   Weeks

$$z = \frac{50 - 41}{.85} = 10.59$$

100%

0   10.59   z scale

41   50   Weeks
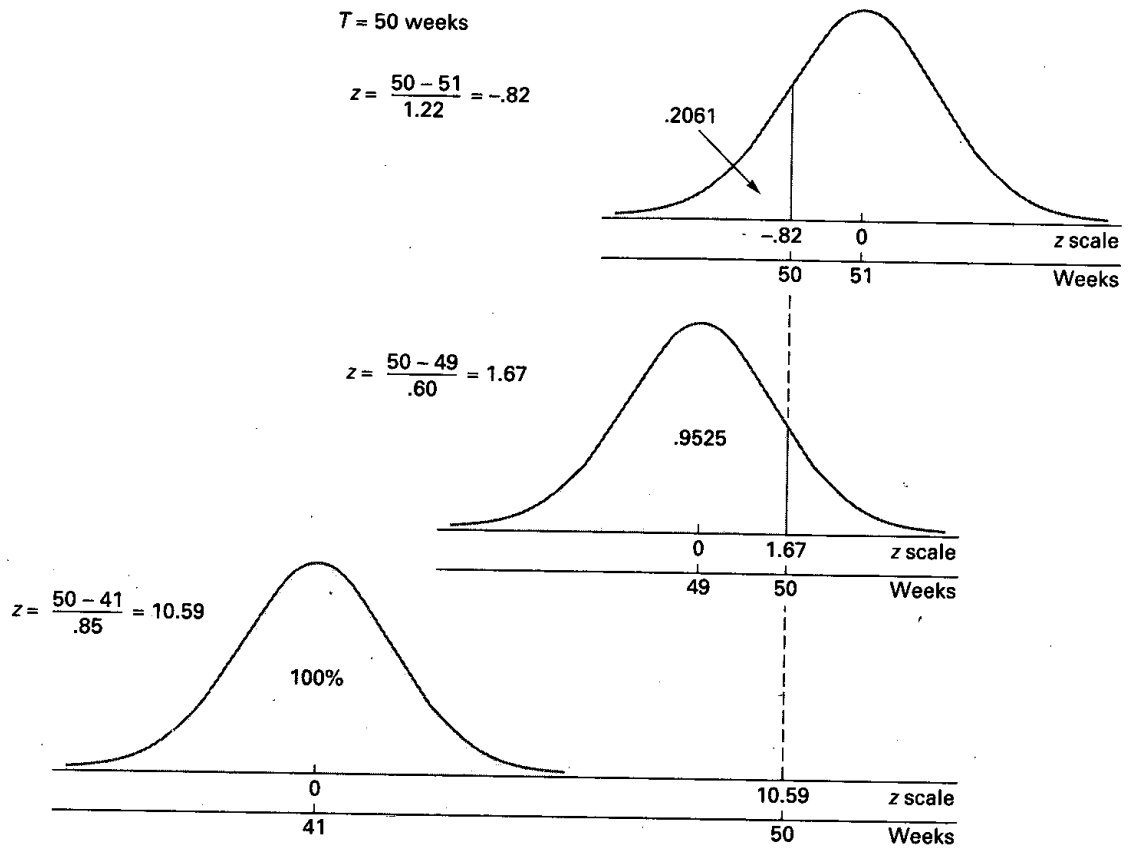
Problem 4

Indirect costs for a project are $12,000 per week for as long as the project lasts. The project manager has supplied the cost and time information shown. Use the information to

a. Determine an optimum crashing plan.

b. Graph the total costs for the plan.

Activity	Crashing Potential (weeks)	Cost per Week to Crash
a	3	$11,000
b	3	3,000 first week, $4,000 others
c	2	6,000
d	1	1,000
e	3	6,000
f	1	2,000

10 weeks
a

14 weeks
b

13 weeks
c

6 weeks
d

15 weeks
e
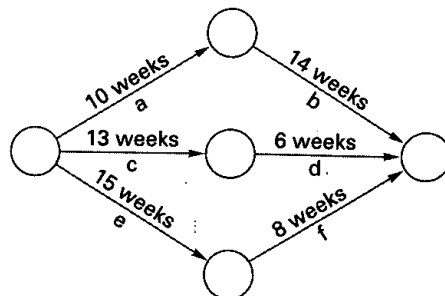
8 weeks
f

Solution

a. (1) Compute path lengths and identify the critical path:

Path	Duration (weeks)
a-b.............	24 (critical path)
c-d.............	19
e-f.............	23

(2) Rank critical activities according to crash costs:

Activity	Cost per Week to Crash
b.............	$ 3,000
a.............	11,000

Activity b should be shortened one week since it has the lower crashing cost. This would reduce indirect costs by $12,000 at a cost of $3,000, for a net savings of $9,000. At this point, paths a-b and e-f would both have a length of 23 weeks, so both would be critical.

(3) Rank activities by crashing costs on the two critical paths:

Path	Activity	Cost per Week to Crash
a-b.........	b	$ 4,000
	a	11,000
e-f.........	e	6,000
	f	2,000

Choose one activity (the least costly) on each path to crash: b on a-b and f on e-f, for a total cost of $4,000 + $2,000 = $6,000 and a net savings of $12,000 − $6,000 = $6,000.

(4) Check to see which path(s) might be critical: a-b and e-f would be 22 weeks in length, and c-d would still be 19 weeks.

(5) Rank activities on the critical paths:

Path	Activity	Cost per Week to Crash
a-b.........	b	$ 4,000
	a	11,000
e-f.........	e	6,000
	f	(no further crashing possible)

Crash b on path a-b and e on e-f for a cost of $4,000 + $6,000 = $10,000, for a net savings of $12,000 − $10,000 = $2,000.

(6) At this point, no further improvement is possible: paths a-b and e-f would be 21 weeks in length, and one activity from each path would have to be shortened. This would mean activity a at $11,000 and e at $6,000 for a total of $17,000, which exceeds the $12,000 potential savings in indirect costs.

b. The following table summarizes the results, showing the length of the project after crashing *n* weeks:

Path	*n* = 0	1	2	3
a-b	24	23	22	21
c-d	19	19	19	19
e-f......................	23	23	22	21
Activity crashed...........		b	b,f	b,e
Crashing costs ($000)		3	6	10

A summary of costs for the preceding schedule would look like this:

Project Length	Cumulative Weeks Shortened	Cumulative Crashing Costs ($000)	Indirect Costs ($000)	Total Costs ($000)
24	0	0	24(12) = 288	288
23	1	3	23(12) = 276	279
22	2	3 + 6 = 9	22(12) = 264	273
21	3	9 + 10 = 19	21(12) = 252	271
20	4	19 + 17 = 36	20(12) = 240	276

The graph of total costs is as follows.
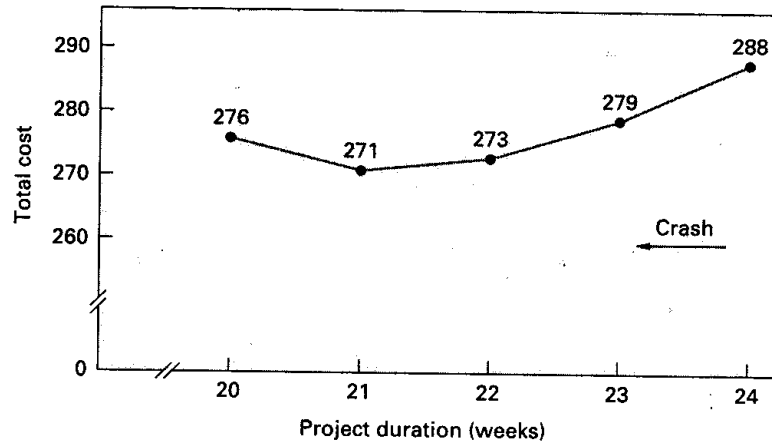
DISCUSSION AND REVIEW QUESTIONS

1. A project manager may need two skill sets—those of a manager and those of a leader. Explain.
2. Explain the term *project champion* and list some ways to keep a champion involved with the project.
3. List the steps in risk management.
4. Give some examples of ethical issues that may arise on projects. What can a project manager do to minimize such issues?
5. What are the key advantages of using project management software?
6. What is a work breakdown structure, and how is it useful for project planning?
7. Identify the term being described for each of the following:
   a. A sequence of activities in a project.
   b. The longest time sequence of activities in a project.
   c. Used when two activities have the same starting and finishing points.
   d. The difference in time length of any path and the critical path.
   e. The statistical distribution used to describe variability of an activity time.
   f. The statistical distribution used to describe path variability.
   g. Shortening an activity by allocating additional resources.
8. List the main advantages of PERT. List the main limitations.
9. Why might a probabilistic estimate of a project's completion time based solely on the variance of the *critical path* be misleading? Under what circumstances would it be acceptable?
10. Define each of these terms, and indicate how each is determined.
    a. Expected activity time.
    b. Variance of an activity time.
    c. Standard deviation of a path's time.

11. Why might a person wish to be involved with a critical path activity? What are some of the reasons one might have for not wanting this association?

12. What are some of the potential benefits of working on a special project in one's firm? What are some of the risks?

13. What are some aspects of the project manager's job that make it more demanding than the job of a manager working in a more routine organizational framework?

14. What is the main benefit of a project organization over more traditional forms of operations management for project work?

1. What trade-offs are associated with time and cost estimates for a proposed project?

2. Who needs to be involved in assessing the cost of a project?

3. Name and explain briefly two different ways that technology has had an impact on project management.

Project management techniques have been used successfully for a wide variety of efforts, including the many NASA space missions, huge construction projects, implementation of major systems such as ERP, production of movies, development of new products and services, theatrical productions, and much more. Why not use them for managing the operations function of any business?

1. For each of the following network diagrams, determine both the critical path and the expected project duration. The numbers on the arrows represent expected activity times.

   a. AOA diagram
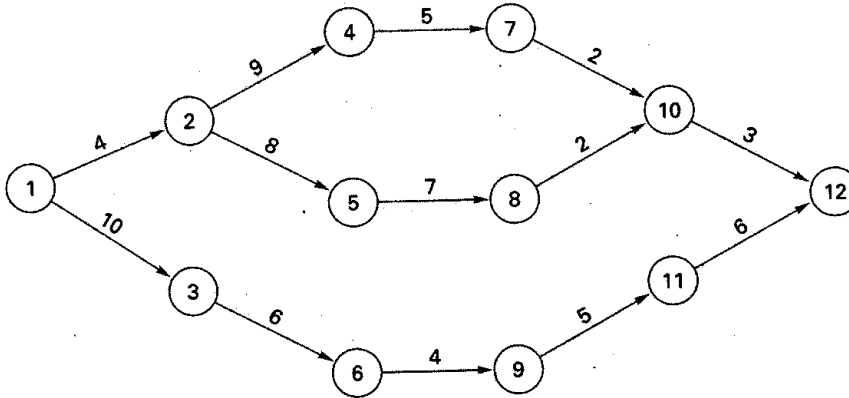
   b. AON diagram
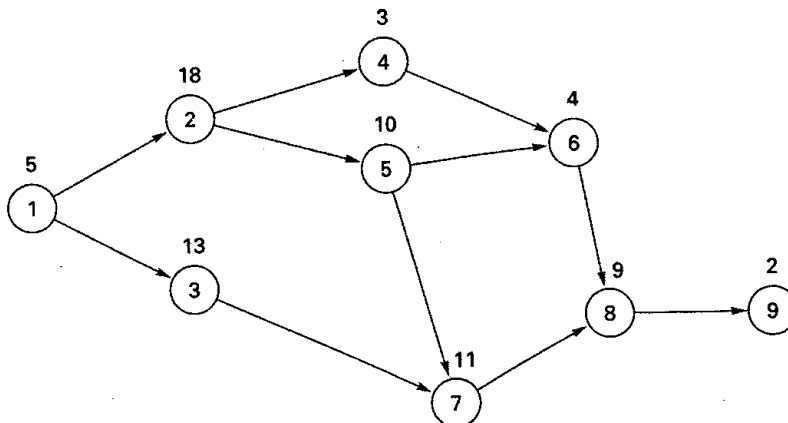
c. AOA diagram
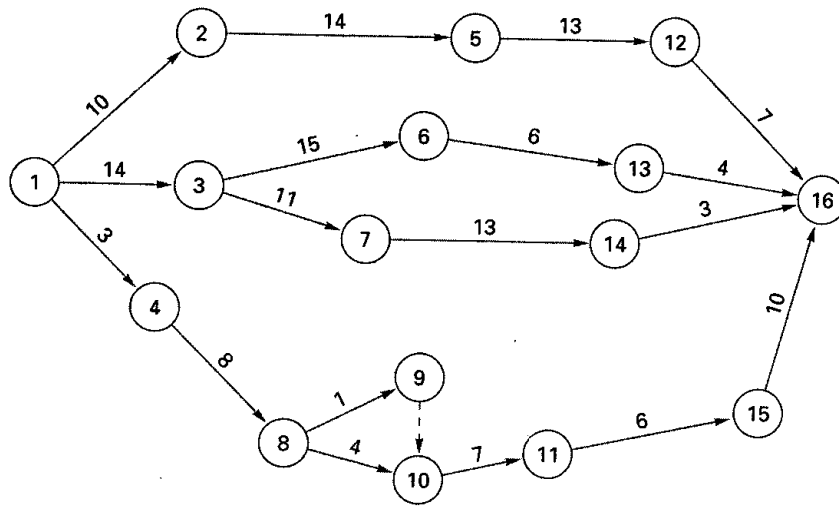
2. Chris received new word processing software for her birthday. She also received a check, with which she intends to purchase a new computer. Chris's college instructor assigned a paper due next week. Chris decided that she will prepare the paper on the new computer. She made a list of the activities she will need to do and their estimated times.
   a. Arrange the activities into two logical sequences.
   b. (1) Construct an AOA network diagram.
      (2) Construct an AON diagram.
   c. Determine the critical path and the expected duration time.
   d. What are some possible reasons for the project to take longer than the expected duration?

Estimated Time (hrs.)	Activity (abbreviation)
0.8	Install software (Inst)
0.4	Outline the paper (Out)
0.2	Submit paper to instructor (Sub)
0.6	Choose a topic (Ch)
0.5	Use grammar-checking routine and make corrections (Ck)
3.0	Write the paper using the word-processing software (Write)
2.0	Shop for a new computer (Sh)
1.0	Select and purchase computer (Sel)
2.0	Library research on chosen topic (Lib)

3. The information in the table pertains to a project that is about to commence. As the project manager, which activities would you be concerned with in terms of timely project completion? Explain.

Activity	Immediate Predecessor	Estimated Time (days)
A	—	15
B	A	12
C	B	6
D	B	5
E	C	3
F	—	8
G	F	8
H	F	9
I	G	7
J	H	14
K	J	6
End	D, E, I, K	

4. *a.* Construct an activity-on-arrow precedence diagram for each of the following cases. Note that each case requires the use of a dummy activity.

  b. Construct an AON diagram for each case.

(1) Activity	Immediate Predecessor	(2) Activity	Immediate Predecessor
A	—	J	—
B	—	K	—
C	—	L	J
D	A	M	L
E	B	N	J
F	B	P	N
G	C	Q	—
H	F	R	K
I	F, G	S	Q
K	D, E	V	R, S, T
End	H, I, K	T	Q
		W	T
		End	M, P, V, W

5. For each of the problems listed, determine the following quantities for each activity: the earliest start time, latest start time, earliest finish time, latest finish time, and slack. List the critical activities, and determine the expected duration of the project.

  a. Problem 1*a.*

  b. Problem 1*b.*

  c. Problem 3.

6. Reconsider the network diagram of Problem 1*a.* Suppose that after 12 weeks, activities 1-2, 1-3, and 2-4 have been finished; activity 2-5 is 75 percent finished; and activity 3-6 is half finished. How many weeks after the original start time should the project be finished?

7. Three recent college graduates have formed a partnership and have opened an advertising firm. Their first project consists of activities listed in the following table.

  a. Draw the precedence diagram.

  b. What is the probability that the project can be completed in 24 days or less? In 21 days or less?

  c. Suppose it is now the end of the seventh day and that activities A and B have been completed while activity D is 50 percent completed. Time estimates for the completion of activity D are 5, 6, and 7. Activities C and H are ready to begin. Determine the probability of finishing the project by day 24 and the probability of finishing by day 21.

Activity	Immediate Predecessor	TIME IN DAYS		
		Optimistic	Most Likely	Pessimistic
A	—	5	6	7
B	—	8	8	11
C	A	6	8	11
D	—	9	12	15
E	C	5	6	9
F	D	5	6	7
G	F	2	3	7
H	B	4	4	5
I	H	5	7	8
End	E, G, I			

8. The new director of special events at a large university has decided to completely revamp graduation ceremonies. Toward that end, a PERT chart of the major activities has been developed. The chart has five paths with expected completion times and variances as shown in the table. Graduation day is 16 weeks from now. Assuming the project begins now, what is the probability that the project will be completed before:

    a. Graduation time?
    b. The end of week 15?
    c. The end of week 13?

Path	Expected Duration (weeks)	Variance
A	10	1.21
B	8	2.00
C	12	1.00
D	15	2.89
E	14	1.44

9. What is the probability that the following project will take more than 10 weeks to complete if the activity means and standard deviations are as shown below?

Activity	Mean	Standard Deviation
1-2	5	1.3
2-3	4	1.0
1-3	8	1.6
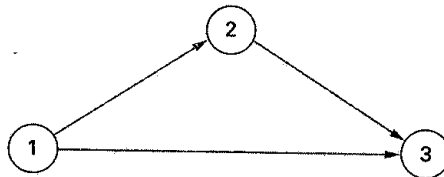
10. The project described in the following table has just begun. It is scheduled to be completed in 11 weeks.
    a. If you were the manager of this project, would you be concerned? Explain.
    b. If there is a penalty of $5,000 a week for each week the project is late, what is the probability of incurring a penalty of at least $5,000?

Activity	Estimated Time (weeks)	Standard Deviation (wks.)
1-2	4	0.70
2-4	6	0.90
1-3	3	0.62
3-4	9	1.90

11. The following precedence diagram reflects three time estimates for each activity. Determine:
    a. The expected completion time for each path and its variance.
    b. The probability that the project will require more than 49 weeks.
    c. The probability that the project can be completed in 46 weeks or less.
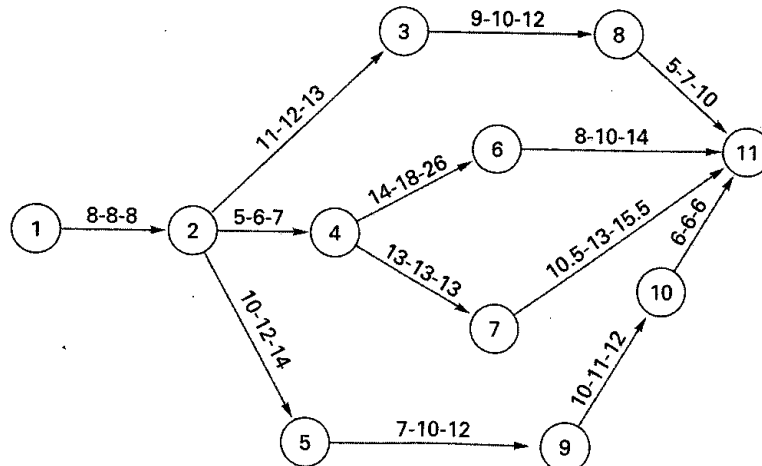
12. A project manager has compiled a list of major activities that will be required to install a computer information system in her firm. The list includes estimated completion times for activities and precedence relationships.

Activity	Immediate Predecessor	Estimated Times (weeks)
A	—	2-4-6
D	A	6-8-10
E	D	7-9-12
H	E	2-3-5
F	A	3-4-8
G	F	5-7-9
B	—	2-2-3
I	B	2-3-6
J	I	3-4-5
K	J	4-5-8
C	—	5-8-12
M	C	1-1-1
N	M	6-7-11
O	N	8-9-13
End	H, G, K, O	

If the project is finished within 26 weeks of its start, the project manager will receive a bonus of $1,000; and if the project is finished within 27 weeks of its start, the bonus will be $500. Find the probability of each bonus.

13. *a.* Construct an activity-on-node diagram for the set of activities listed in Problem 3.
    b. Construct an activity-on-node diagram for the set of activities listed in Problem 12.

14. The project manager of a task force planning the construction of a domed stadium had hoped to be able to complete construction prior to the start of the next college football season. After reviewing construction time estimates, it now appears that a certain amount of crashing will be needed to ensure project completion before the season opener. Given the following time and cost estimates, determine a minimum-cost crashing schedule that will shave five weeks off the project length. *Note:* No activity can be crashed more than two weeks.

Activity	Immediate Predecessor	Normal Time (weeks)	CRASHING COSTS	
			First Week	Second Week
A	—	12	$15,000	$20,000
B	A	14	10,000	10,000
C	—	10	5,000	5,000
D	C	17	20,000	21,000
E	C	18	16,000	18,000
F	C	12	12,000	15,000
G	D	15	24,000	24,000
H	E	8	—	—
I	F	7	30,000	—
J	I	12	25,000	25,000
K	B	9	10,000	10,000
M	G	3	—	—
N	H	11	40,000	—
P	H, J	8	20,000	20,000
End	K, I, M, N, P			

15. A construction project has indirect costs totaling $40,000 per week. Major activities in the project and their expected times are shown in this precedence diagram:
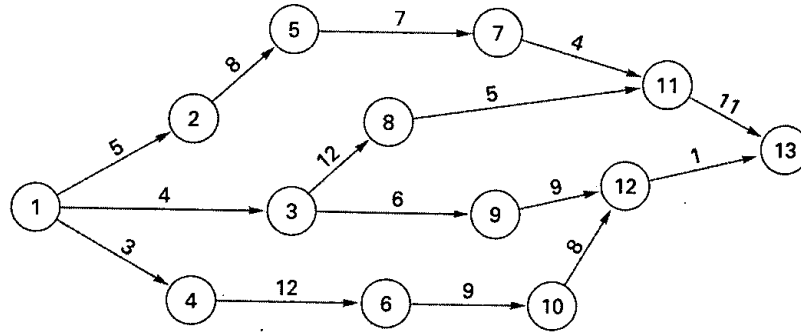
Crashing costs for each activity are

CRASHING COSTS ($000)

Activity	First Week	Second Week	Third Week
1-2	$18	$22	$—
2-5	24	25	25
5-7	30	30	35
7-11	15	20	—
11-13	30	33	36
1-3	12	24	26
3-8	—	—	—
8-11	40	40	40
3-9	3	10	12
9-12	2	7	10
12-13	26	—	—
1-4	10	15	25
4-6	8	13	—
6-10	5	12	—
10-12	14	15	—

a. Determine the optimum time–cost crashing plan.
b. Plot the total-cost curve that describes the least expensive crashing schedule that will reduce the project length by six weeks.

16. Chuck's Custom Boats (CCB) builds luxury yachts to customer order. CCB has landed a contract with a mysterious New York lawyer (Mr. T). Relevant data are shown below. The complication is that Mr. T wants delivery in 32 weeks or he will impose a penalty of $375 for each week his yacht is late. *Note*: No activity can be crashed more than two weeks.

Activity	Immediate Predecessor	Normal Time (weeks)	CRASHING COSTS	
			1st Week	2nd Week
K	—	9	$410	$415
L	K	7	125	—
N	K	5	45	45
M	L	4	300	350
J	N	6	50	—
Q	J, M	5	200	225
P	Q	8	—	—
Y	Q	7	85	90
Z	P	6	90	—
End	Y, Z			

Develop a crashing schedule.

The Case of the Mexican Crazy Quilt   CASE

"The mission of the project which you will head is to get our new Mexican subsidiary ready for take-over by Mexican managers. My hope is that you will be able to do this in about two years," explained Robert Linderman, president of Linderman Industries, Inc., to Carl Conway, newly appointed manager for "Operation Mexicano." Conway had been hired specifically for this assignment because of his experience in managing large defense projects in the aerospace industry.

"The first thing that I will have to do is put a project team together," said Conway. "I imagine that you have in mind my drawing people from the functional divisions."

"Yes, I have already sent memoranda to the division managers informing them that you will be asking for some of their key people to work under you for about two years," said Linderman. "In addition, I have advised them to be prepared to process work orders from Operation Mexicano with the personnel and equipment of their organizations. Later on in the project's life, you will begin to get Mexican personnel, both managers and technicians, in to your organization. These people will have Mexican supervisors, but until the mission is accomplished, they also will report to you. I will have to admit that you are going to have some complex authority relationships, especially as you personally will be responsible to the president of the subsidiary, Felix Delgado, as well as to me."

Conway began to make his plans for the project team. The plant building was available and empty in Mexico City, and it was important to get equipment purchased and installed as soon as possible. A plant layout would have to be prepared, but before that could be done there would have to be a manufacturing plan. Therefore, he needed to recruit an industrial engineer, a production planner, and an equipment buyer. They, in turn, would have to build their own staffs.

He made an appointment with Sam Sargis, corporate manager of industrial engineering. "I have had a preliminary talk with Bob Cates about his joining Operation Mexicano, and he is quite interested," Carl said. "Will you release him to me?"

"Why, I'm grooming Cates to take over my job when I retire," replied Sargis. "He is my best man. Let me pick someone else for you, or better still, you just tell me what industrial engineering work you want done, and I will have it done for you."

"Sorry, I want Cates," said Carl firmly. "And besides, you are not due to retire for five years. This will be good experience for him."

For production planning, Carl had in mind Bert Mill, an older man with extensive experience in managing production operations, but Mill rejected his offer. "I talked it over with my wife,"

he said, "and we feel that at my age I shouldn't take a chance on not having a job to come back to when Operation Mexicano is finished."

Carl next talked to Emil Banowetz, who was assistant to Jim Burke, the vice president for manufacturing, and Banowetz decided that he would like to join the project team. However, Burke told Conway that if Banowetz were forcibly taken away from him, he would give Mr. Linderman his resignation, so Carl decided to back down. He finally accepted a man that Burke recommended.

Filling the equipment buyer's slot was easy. The director of procurement phoned Carl and said that a senior buyer, Humberto Guzman, had requested permission to ask for the assignment, and that he strongly recommended him. Guzman has been purchasing agent for a large mining company in Mexico for about 10 years.

Carl had about the same experiences in getting the people he wanted for the functions of engineering, quality control, cost, marketing, and advertising as he did for the first three positions; in other words, he won some confrontations with the division managers and lost some. For personnel, he got Dr. Juan Perez, who was slated to be personnel director of the subsidiary company, to affiliate temporarily with the project team.

The first brush that Project Mexicano had in getting a functional division to do work for it came when Carl's engineering man, Frank Fong, reported to him that the engineering vice president, who was formerly Fong's boss, refused to authorize top priority to the changing of dimensions in the production drawings to the metric system. Carl had to take this issue to Linderman, who ruled in his favor. The defeated vice president, of course, did not take kindly to the decision.

The next incident revolved around Carl's desire to have a pilot run of products made with metric measurements for shipment to Mexico. The purpose was to test the market acceptance of the Linderman articles. Jim Burke stated flatly that there was no way that his production workers could be trained to work with metric drawings. Carl quickly saw that this was an issue that he was not going to win, so he had his buyer, Guzman, work with the newly appointed manufacturing manager for the subsidiary in getting a run of the products subcontracted in Mexico City.

Bob Cates made a special trip from Mexico City to present Carl with an interesting problem. The Mexican industrial engineer, whom Bob was supposed to be training, had his own ideas about plant layout. When they differed from Bob's as they usually did, he would take his complaint directly to

(concluded)

Felix Delgado, the president of the Mexican subsidiary. Because Delgado's competence was primarily in finance, he would not know how to decide the argument and would simply table it. Carl took examples of some of the disagreements to Bob's former boss, Sam Sargis, who quite unexpectedly ruled against Bob's proposed methods. Carl saw that there was bad feeling by Sargis against Bob for leaving his department, which boded ill for Bob's return. To solve the immediate problem, however, Carl asked Dr. Perez to try to reconcile the situation in Mexico City.

Despite these problems, and many more of a similar nature, Project Mexicano was successful, and the transition to Mexican management was made in just a little over two years. By a curious twist, through Dr. Perez's intercession Felix Delgado became very impressed by Bob Cates and convinced him to accept the job of director of industrial engineering for the Mexican company. Humberto Guzman also stayed on to head the procurement operation.

Other members of the project team were not so fortunate. Linderman Industries was laying off personnel when the project ended, and only the project production man was able to get a job in the company at as high a level as the one he had when he joined the team. The cost expert elected to leave Linderman because he said the glamour of Project Mexicano had spoiled him for any routine job.

Carl Conway had a difficult decision of his own to make. Robert Linderman said that he was extremely pleased with his

performance and that something good would open up in the company for him soon. In the meantime, there was a staff assignment available for him. Carl had seen enough project managers in the aerospace industry who had figuratively rotted on staff assignments when their projects were completed to be somewhat wary.

Questions

1. Was Linderman Industries' adoption of project organization an appropriate one for getting the Mexican subsidiary started?
2. In consideration of Robert Linderman's letting the division managers know that the project manager would be asking for some of their key people, why would Conway have any difficulty in getting the ones he wanted?
3. Would you expect that many people would turn down a chance to join a project organization, as Bert Mill did?
4. Why would Conway take his problem with the engineering vice president to Linderman and have it resolved in his favor, yet back down in two disputes with the manufacturing vice president?
5. What could Linderman Industries have done to assure good jobs for the people coming off Project Mexicano, including Carl Conway, the project manager?

Time, Please   CASE

B. "Smitty" Smith is a project manager for a large consumer electronics corporation. Although she has been with the company only four years, she has demonstrated an uncanny ability to bring projects in on time, meet technical specifications, and be close to budget.

Her latest assignment is a project that will involve merging two existing technologies. She and her team have almost finished developing the proposal that will be presented to a management committee for approval. All that remains to be done is to develop a time estimate for the project. The team has to construct a network diagram for the project. It has three paths. The expected durations and standard deviations for the paths are listed in the following table.

Path	Expected Duration (weeks)	Standard Deviation
A	10	4
B	14	2
C	13	2

What project durations (in weeks) should Smitty include in the proposal for these risks of not delivering the project on time: 5 percent, 10 percent, 15 percent? What are the pros and cons of quoting project times aggressively? Conservatively?

Angus, Robert B., Norman A. Gundersen, and Thomas P. Cullinane. *Planning, Performing, and Controlling Projects: Principles and Applications.* 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2000.

Chapman, Chris, and Stephen Ward. *Project Risk Management.* New York: John Wiley & Sons, 1997.

Cleland, David I. *Project Managers' Portable Handbook.* New York: McGraw-Hill, 2000.

DeWeaver, Mary Feeherry, and Lori Ciprian Gillespie. *Real World Project Management.* New York: Quality Resources, 1997.

Ghattas, R. G., and Sandra L. McKee. *Practical Project Management.* Upper Saddle River, NJ: Prentice Hall, 2001.

Gido, Jack, and James P. Clements. *Successful Project Management.* Cincinnati: South-Western, 1999.

Goldratt, Eliyahu M. *Critical Chain.* Great Barrington, MA: North River Press, 1997.

Graham, Robert J., and Randall L. Englund. *Creating an Environment for Successful Projects: The Quest to Manage Project Management.* San Francisco: Jossey-Bass, 1997.

Gray, Clifford F., and Erik W. Larson. *Project Management: The Managerial Process.* 2nd ed. New York: McGraw-Hill/Irvin, 2003.

Kanabar, Vijay. *Project Risk Management: A Step-by-Step Guide to Reducing Project Risk.* Boston: Copley Publishing Group, 1997.

Kerzner, Harold. *Project Management: A Systems Approach for Planning, Scheduling, and Controlling.* 7th ed. New York: Wiley, 2001.

Larson, Melissa. "Manage Your Projects before They Manage You." *Quality,* September 1997, pp. 64–67.

Lientz, Bennet P., and Kathryn P. Rea. *Breakthrough Project Management.* New York: Academic Press, 1999.

Mantel, Samuel, Jack R. Meredith, Scott M. Shafer, and Margaret Sutton. *Project Management in Practice.* New York: Wiley, 2001.

Matta, Nadim F., and Ronald N. Ashkenas. "Why Good Projects Fail Anyway." *Harvard Business Review,* September 2003, pp. 109–114.

Project Management Institute Standards Committee. *A Guide to the Project Management Body of Knowledge.* Upper Darby, PA: Project Management Institute Communications, 1996.

Smith-Daniels, Dwight. "Teaching Project Management to MBAs: The Means to How Many Ends?" *Decision Line,* May 1997, pp. 11–13.

Stevenson, William J. *Introduction to Management Science.* 3rd ed. New York: McGraw-Hill/Irwin, 1998.

SELECTED BIBLIOGRAPHY AND FURTHER READING

Waiting Lines and Simulation

The chapter in this section covers the analysis of waiting lines and simulation. The occurrence of waiting lines in service systems indicates an imbalance between demand and capacity. Analysis can help managers reduce the impact of waiting lines on system cost and effectiveness.

The chapter supplement covers simulation, a descriptive tool useful for studying system behavior.

CHAPTER 18

Waiting Lines

LEARNING OBJECTIVES

After completing this chapter, you should be able to:

1. Explain why waiting lines form in systems that are underloaded.
2. Identify the goal of queuing (waiting-line) analysis.
3. List the measures of system performance that are used in queuing.
4. Discuss the assumptions of the basic queuing models presented.
5. Solve typical problems.

CHAPTER OUTLINE

The mission of Walt Disney theme parks is to "make people happy," and the folks at Disney World in Orlando are masters at doing that. They realize that waiting in lines at attractions does not add to the enjoyment of their customers. They also realize that customers waiting in lines are not generating the revenue that they would if they visited restaurants and souvenir shops. Hence, they have several reasons for wanting to reduce waiting times. Lately they have been using reservation systems at some attractions that allow customers to reserve visit times instead of having to wait in line. This is a win-win solution: Customers are happier because they don't have to wait in line, and the park's potential for additional revenue is increased.

Waiting lines abound in all sorts of service systems. And they are *non-valued-added* occurrences. For customers, having to wait for service can range from being acceptable (usually short waits), to being annoying (longer waits), to being a matter of life and death (e.g., in emergencies). For businesses, the costs of waiting come from lower productivity and competitive disadvantage. For society, the costs are wasted resources (e.g., fuel consumption of cars stuck in traffic) and reduced quality of life. Hence, it is important for system designers and managers of existing service systems to fully appreciate the impact of waiting lines.

Designers must weigh the cost of providing a given level of service capacity against the potential (implicit) cost of having customers wait for service. This planning and analysis of service capacity frequently lends itself to **queuing theory,** which is a mathematical approach to the analysis of waiting lines.

The foundation of modern queuing theory is based on studies about automatic dialing equipment made in the early part of the twentieth century by Danish telephone engineer A. K. Erlang. Prior to World War II, very few attempts were made to apply queuing theory to business problems. Since that time, queuing theory has been applied to a wide range of problems.

www.disney.com

queuing theory Mathematical approach to the analysis of waiting lines.

The mathematics of queuing can be complex; for that reason, the emphasis here will not be on the mathematics but the concepts that underlie the use of queuing in analyzing waiting-line problems. We shall rely on the use of formulas and tables for analysis.

Waiting lines are commonly found wherever customers arrive *randomly* for services. Some examples of waiting lines we encounter in our daily lives include the lines at super-market checkouts, fast-food restaurants, airport ticket counters, theaters, post offices, and toll booths. In many situations, the "customers" are not people but orders waiting to be filled, trucks waiting to be unloaded, jobs waiting to be processed, or equipment awaiting re-pairs. Still other examples include ships waiting to dock, planes waiting to land, and cars waiting at a stop sign.

One reason that queuing analysis is important is that customers regard waiting negatively. Customers may tend to associate this with poor service quality, especially if the wait is long. Similarly, in an organizational setting, having work or employees wait is the sort of waste that workers in JIT systems strive to reduce.

The discussion of queuing begins with an examination of what is perhaps the most funda-mental issue in waiting-line theory: Why is there waiting?

WHY IS THERE WAITING?

Many people are surprised to learn that waiting lines tend to form even though a system is ba-sically underloaded. For example, a fast-food restaurant may have the capacity to handle an av-erage of 200 orders per hour and yet experience waiting lines even though the average number of orders is only 150 per hour. The key word is *average*. In reality, customers arrive at random intervals rather than at evenly spaced intervals, and some orders take longer to fill than others. In other words, both arrivals and service times exhibit a high degree of variability. As a result, the system at times becomes temporarily overloaded, giving rise to waiting lines; at other times, the system is idle because there are no customers. It follows that in systems where vari-ability is minimal or nonexistent (e.g., because arrivals can be scheduled and service time is constant), waiting lines do not ordinarily form.

MANAGERIAL IMPLICATIONS OF WAITING LINES

Managers have a number of very good reasons to be concerned with waiting lines. Chief among those reasons are the following:

1. The cost to provide waiting space.
2. A possible loss of business should customers leave the line before being served or refuse to wait at all.
3. A possible loss of goodwill.
4. A possible reduction in customer satisfaction.
5. The resulting congestion that may disrupt other business operations and/or customers.

GOAL OF WAITING-LINE ANALYSIS

The goal of queuing is essentially to minimize total costs. There are two basic categories of cost in a queuing situation: those associated with customers waiting for service and those as-sociated with capacity. Capacity costs are the costs of maintaining the ability to provide ser-vice. Examples include the number of bays at a car wash, the number of checkouts at a supermarket, the number of repair people to handle equipment breakdowns, and the number of lanes on a highway. When a service facility is idle, capacity is lost since it cannot be stored. The costs of customer waiting include the salaries paid to employees while they wait for ser-vice (mechanics waiting for tools, the drivers of trucks waiting to unload), the cost of the space

Cost

Total cost

Cost of service capacity

Cost of customers waiting

0

Optimum

Service capacity

for waiting (size of doctor's waiting room, length of driveway at a car wash, fuel consumed by planes waiting to land), and any loss of business due to customers refusing to wait and possibly going elsewhere in the future.

A practical difficulty frequently encountered is pinning down the cost of customer waiting time, especially since major portions of that cost are not a part of accounting data. One approach often used is to treat waiting times or line lengths as a policy variable: A manager simply specifies an acceptable level of waiting and directs that capacity be established to achieve that level.

The goal of queuing analysis is to balance the cost of providing a level of service capacity with the cost of customers waiting for service. Figure 18.1 illustrates this concept. Note that as capacity increases, its cost increases. For simplicity, the increase is shown as a linear relationship. Although a step function is often more appropriate, use of a straight line does not significantly distort the picture. As capacity increases, the number of customers waiting and the time they wait tend to decrease, thereby decreasing waiting costs. As is typical in trade-off relationships, total costs can be represented as a U-shaped curve. The goal of analysis is to identify a level of service capacity that will minimize total cost. (Unlike the situation in the inventory EOQ model, the minimum point on the total cost curve is *not* usually where the two cost lines intersect.)

In situations where those waiting in line are *external* customers (as opposed to employees), the existence of waiting lines can reflect negatively on an organization's *quality* image. Consequently, some organizations are focusing their attention on providing faster service—speeding up the rate at which service is delivered rather than merely increasing the number of servers. The effect of this is to shift the total cost curve downward if the cost of customer waiting decreases by more than the cost of the faster service.

CHARACTERISTICS OF WAITING LINES

There are numerous queuing models from which an analyst can choose. Naturally, much of the success of the analysis will depend on choosing an appropriate model. Model choice is affected by the characteristics of the system under investigation. The main characteristics are

1. Population source.
2. Number of servers (channels).

A simple queuing system
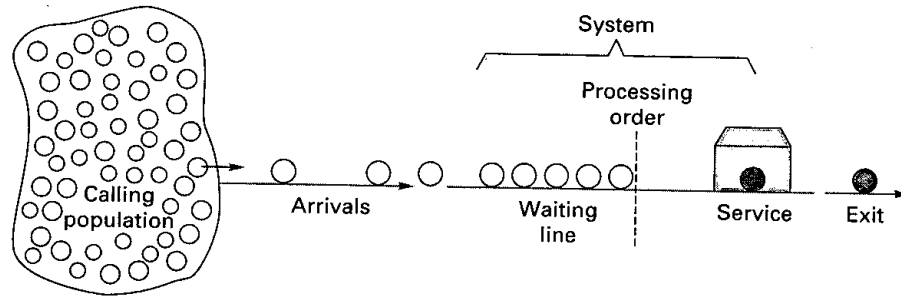
3. Arrival and service patterns.
4. Queue discipline (order of service).

Figure 18.2 depicts a simple queuing system.

Population Source

infinite source Customer arrivals are unrestricted.

finite source The number of potential customers is limited.

The approach to use in analyzing a queuing problem depends on whether the potential number of customers is limited. There are two possibilities: *infinite-source* and *finite-source* populations. In an **infinite-source** situation, the *potential* number of customers greatly exceeds system capacity. Infinite-source situations exist whenever service is *unrestricted*. Examples are supermarkets, drugstores, banks, restaurants, theaters, amusement centers, and toll bridges. Theoretically, large numbers of customers from the "calling population" can request service at any time. When the potential number of customers is limited, a **finite-source** situation exists. An example is the repairman responsible for a certain number of machines in a company. The potential number of machines that might need repairs at any one time cannot exceed the number of machines assigned to the repairer. Similarly, an operator may be responsible for loading and unloading a bank of four machines, a nurse may be responsible for answering patient calls for a 10-bed ward, a secretary may be responsible for taking dictation from three executives, and a company shop may perform repairs as needed on the firm's 20 trucks.

Number of Servers (Channels)

channel A server in a service system.

The capacity of queuing systems is a function of the capacity of each server and the number of servers being used. The terms *server* and **channel** are synonymous, and it is generally assumed that each channel can handle one customer at a time. Systems can be either *single-* or *multiple-channel*. (A group of servers working together as a team, such as a surgical team, is treated as a single-channel system.) Examples of single-channel systems are small grocery stores with one checkout counter, some theaters, single-bay car washes, and drive-in banks with one teller. Multiple-channel systems (those with more than one server) are commonly found in banks, at airline ticket counters, at auto service centers, and at gas stations.

A related distinction is the number of steps or *phases* in a queuing system. For example, at theme parks, people go from one attraction to another. Each attraction constitutes a separate phase where queues can (and usually do) form.

Figure 18.3 illustrates some of the most common queuing systems. Because it would not be possible to cover all of these cases in sufficient detail in the limited amount of space available here, our discussion will focus on *single-phase* systems.

Arrival and Service Patterns

Waiting lines are a direct result of arrival and service variability. They occur because random, highly variable arrival and service patterns cause systems to be temporarily overloaded. In many instances, the variabilities can be described by theoretical distributions. In fact, the most commonly used models assume that arrival and service *rates* can be described by a Poisson distribution or, equivalently, that the interarrival *time* and service *time* can be described by a negative exponential distribution. Figure 18.4 illustrates these distributions.
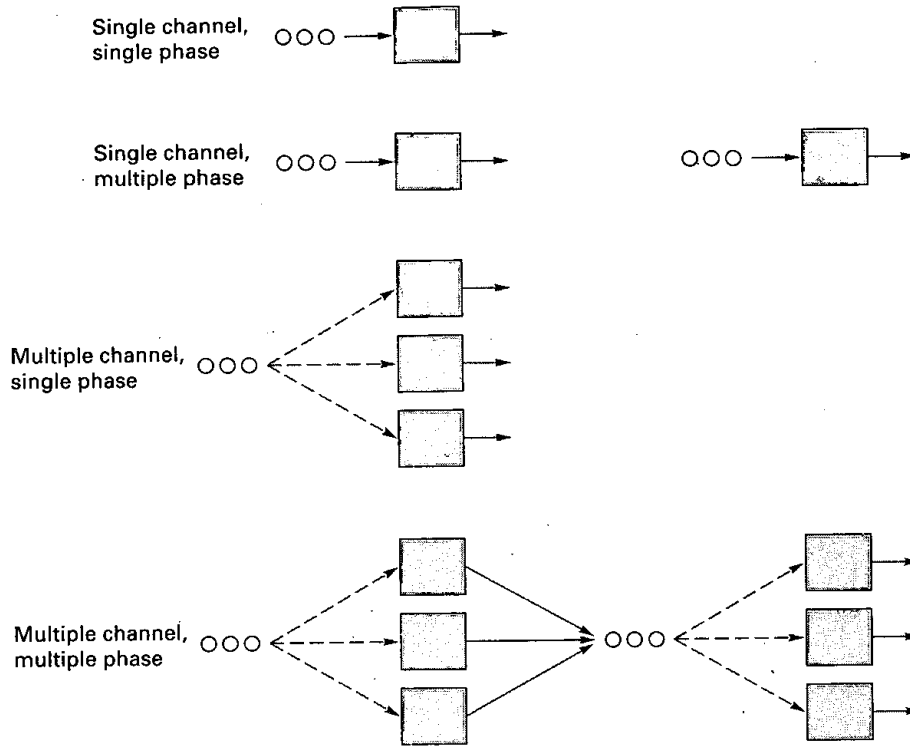
FIGURE 18.3

Four common variations of queuing systems

Single channel, single phase

Single channel, multiple phase

Multiple channel, single phase

Multiple channel, multiple phase
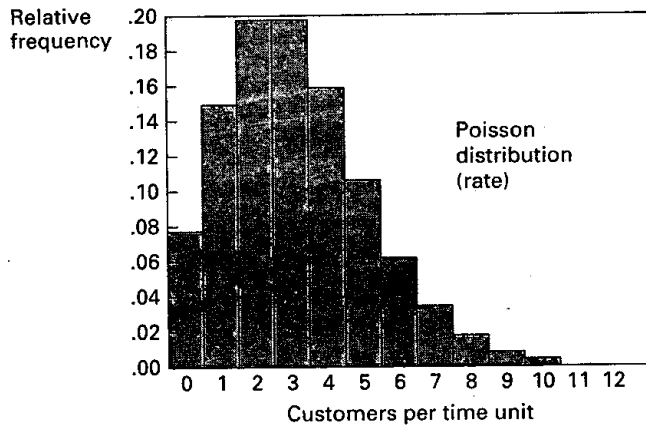
FIGURE 18.4

Poisson and negative exponential distributions

Relative frequency

Poisson distribution (rate)

Customers per time unit

Relative frequency (%)

Exponential distribution (time)
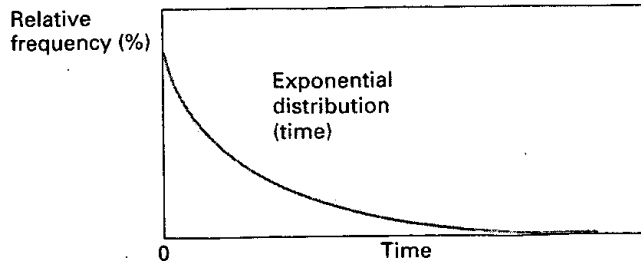
Time

The Poisson distribution often provides a reasonably good description of customer arrivals per unit of time (e.g., per hour). Figure 18.5A illustrates how Poisson-distributed arrivals (e.g., accidents) might occur during a three-day period. In some hours, there are three or four arrivals; in other hours one or two arrivals; and in some hours no arrivals.

FIGURE 18.5

Poisson arrivals and exponential service times

A. Customer Arrivals
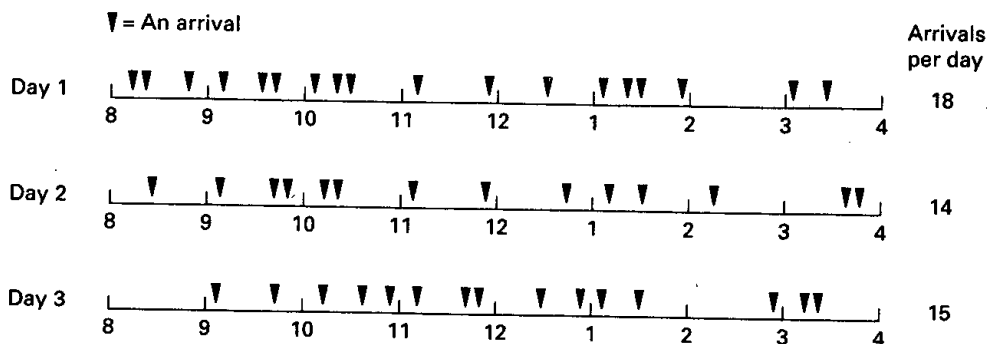
B. Service Times

IOM
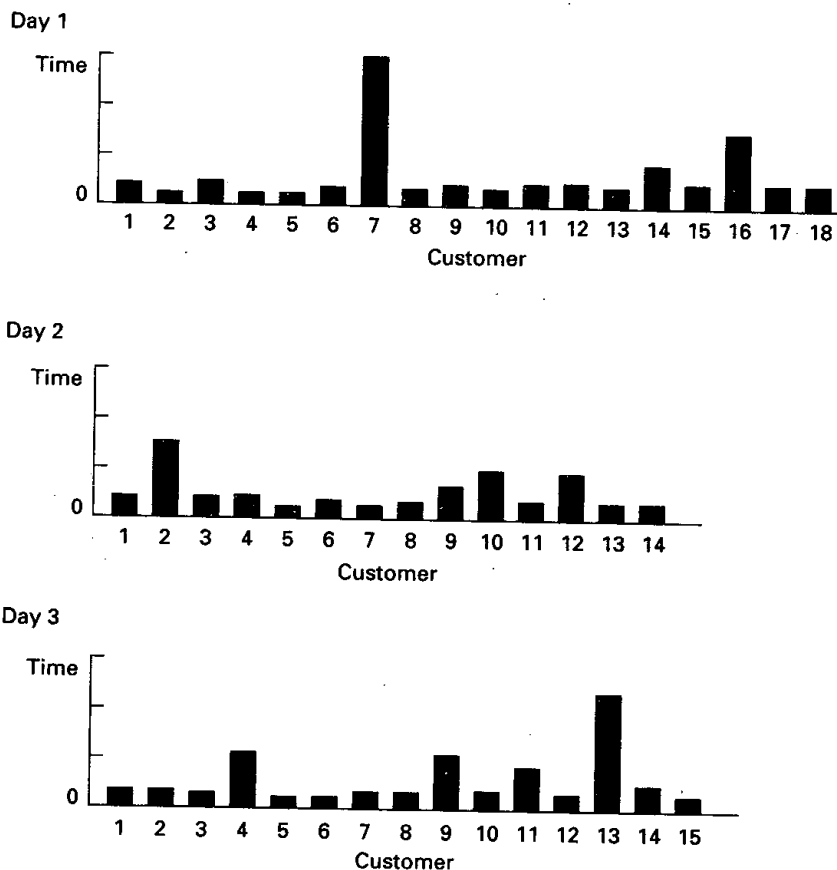
The negative exponential distribution often provides a reasonably good description of customer service times (e.g., first aid care for accident victims). Figure 18.5B illustrates how exponential service times might appear for the customers whose arrivals are illustrated in Figure 18.5A. Note that most service times are very short—some are close to zero—but a few require a relatively long service time. That is typical of a negative exponential distribution.

Waiting lines are most likely to occur when arrivals are bunched or when service times are particularly lengthy, and they are very likely to occur when both factors are present. For instance, note the long service time of customer 7 on day 1, in Figure 18.5B. In Figure 18.5A, the seventh customer arrived just after 10 o'clock and the next two customers arrived shortly after that, making it very likely that a waiting line formed. A similar situation occurred on day 3 with the last three customers: The relatively long service time for customer 13 (Figure 18.5B), and the short time before the next two arrivals (Figure 18.5A, day 3) would create (or increase the length of) a waiting line.
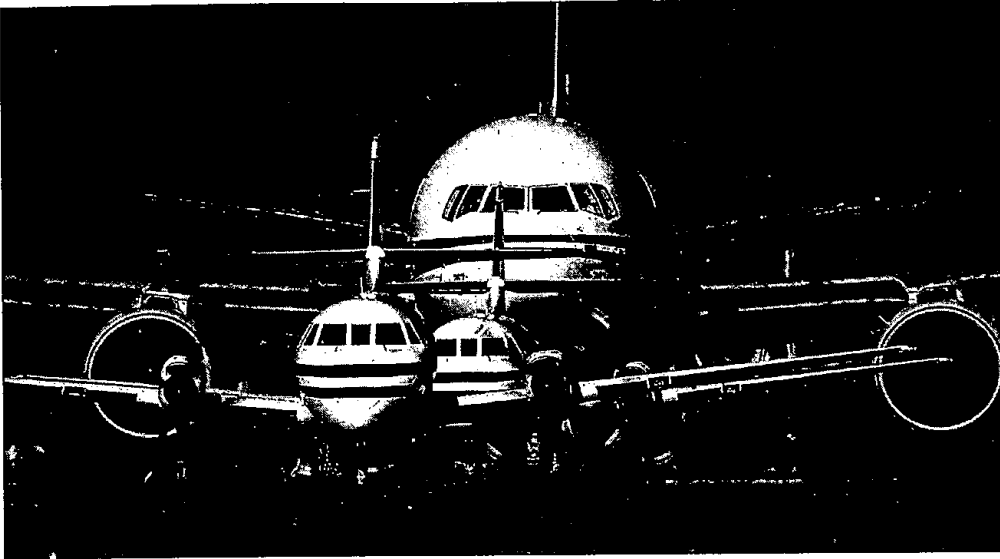
Departing aircraft can face lengthy waiting times due to predictable events, such as heavy demand, or unpredictable circumstances such as bad weather, mechanical problems, and security delays at the airport.

It is interesting to note that the Poisson and negative exponential distributions are alternate ways of presenting the same basic information. That is, if service time is exponential, then the service rate is Poisson. Similarly, if the customer arrival rate is Poisson, then the interarrival time (i.e., the time between arrivals) is exponential. For example, if a service facility can process 12 customers per hour (rate), average service time is five minutes. And if the arrival rate is 10 per hour, then the average time between arrivals is six minutes.

The models described here generally require that arrival and service rates lend themselves to description using a Poisson distribution or, equivalently, that interarrival and service times lend themselves to description using a negative exponential distribution. In practice, it is necessary to verify that these assumptions are met. Sometimes this is done by collecting data and plotting them, although the preferred approach is to use a chi-square goodness-of-fit test for that purpose. A discussion of the chi-square test is beyond the scope of this text, but most basic statistics textbooks cover the topic.

Research has shown that these assumptions are often appropriate for customer arrivals but less likely to be appropriate for service. In situations where the assumptions are not reasonably satisfied, the alternatives would be to (1) develop a more suitable model, (2) search for a better (and usually more complex) existing model, or (3) resort to computer simulation. Each of these alternatives requires more effort or cost than the ones presented here.

The models in this chapter assume customers are patient; that is, that customers enter the waiting line and remain until they are served. Other possibilities are (1) waiting customers grow impatient and leave the line (*reneging*); (2) if there are multiple waiting customers, customers may switch to another line (*jockeying*); or (3) upon arriving, decide the line is too long and, therefore, not to enter the line (*balking*).

Queue Discipline

Queue discipline refers to the order in which customers are processed. All but one of the models to be described shortly assume that service is provided on a *first-come, first-served* basis. This is perhaps the most commonly encountered rule. There is first-come service at banks, stores, theaters, restaurants, four-way stop signs, registration lines, and so on. Examples of systems that do not serve on a first-come basis include hospital emergency rooms, rush orders in a factory, and main frame computer processing of jobs. In these and similar situations, customers do not all represent the same waiting costs; those with the highest costs (e.g., the most seriously ill) are processed first, even though other customers may have arrived earlier.

queue discipline The order in which customers are processed.

FIGURE 18.6

The average number waiting in line and the average time customers wait in line increase exponentially as the system utilization increases
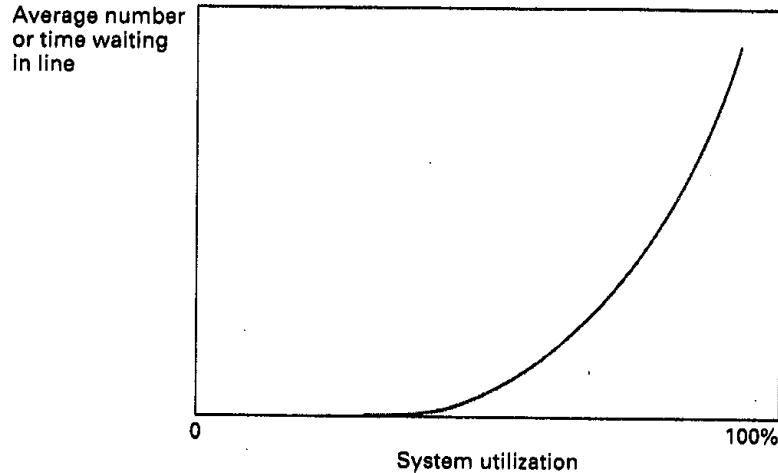
MEASURES OF WAITING LINE PERFORMANCE

The operations manager typically looks at five measures when evaluating existing or proposed service systems:

1. The average number of customers waiting, either in line or in the system.
2. The average time customers wait, either in line or in the system.
3. System utilization, which refers to the percentage of capacity utilized.
4. The implied cost of a given level of capacity and its related waiting line.
5. The probability that an arrival will have to wait for service.

Of these measures, system utilization bears some elaboration. It reflects the extent to which the servers are busy rather than idle. On the surface, it might seem that the operations manager would want to seek 100 percent utilization. However, as Figure 18.6 illustrates, increases in system utilization are achieved at the expense of increases in both the length of the waiting line and the average waiting time. In fact, these values become exceedingly large as utilization approaches 100 percent. The implication is that under normal circumstances, 100 percent utilization is not a realistic goal. Even if it were, 100 percent utilization of service personnel is not good; they need some slack time. Thus, instead, the operations manager should try to achieve a system that minimizes the sum of waiting costs and capacity costs.

QUEUING MODELS: INFINITE-SOURCE

Many queuing models are available for a manager or analyst to choose from. The discussion here includes four of the most basic and most widely used models. The purpose is to provide an exposure to a range of models rather than an extensive coverage of the field. All assume a Poisson arrival rate. Moreover, the models pertain to a system operating under *steady-state* conditions; that is, they assume the average arrival and service rates are stable. The four models described are

1. Single-channel, exponential service time.
2. Single-channel, constant service time.
3. Multiple-channel, exponential service time.
4. Multiple-priority-service, exponential service time.

To facilitate your use of queuing models, Table 18.1 provides a list of the symbols used for the infinite-source models.

Symbol	Represents
λ	Customer arrival rate
μ	Service rate per server
L_q	The average number of customers waiting for service
L_s	The average number of customers in the system (waiting and/or being served)
r	The average number of customers being served
ρ	The system utilization
W_q	The average time customers wait in line
W_s	The average time customers spend in the system (waiting in line and service time)
$1/\mu$	Service time
P_0	The probability of zero units in the system
P_n	The probability of n units in the system
M	The number of servers (channels)
L_{max}	The maximum expected number waiting in line

Basic Relationships

There are certain basic relationships that hold for all infinite-source models. Knowledge of these can be very helpful in deriving desired performance measures, given a few key values. Here are the basic relationships:

Note: The arrival and service rates, represented by λ and μ, must be in the same units (e.g., customers per hour, customers per minute).

System utilization: This reflects the ratio of demand (as measured by the arrival rate) to supply or capacity (as measured by the product of the number of servers, M, and the service rate, μ).

$$\rho = \frac{\lambda}{M\mu} \tag{18-1}$$

The average number of customers being served:

$$r = \frac{\lambda}{\mu} \tag{18-2}$$

The average *number* of customers

Waiting in line for service: L_q [Model dependent. Obtain using a table or formula.]

In the system (line plus being served): $L_s = L_q + r$ $\tag{18-3}$

The average *time* customers are

Waiting in line: $W_q = \dfrac{L_q}{\lambda}$ $\tag{18-4}$

In the system: $W_s = W_q + \dfrac{1}{\mu} = \dfrac{L_s}{\lambda}$ $\tag{18-5}$

All infinite-source models require that system utilization be less than 1.0; the models apply only to underloaded systems.

The average number waiting in line, L_q, is a key value because it is a determinant of some of the other measures of system performance, such as the average number in the system, the average time in line, and the average time in the system. Hence, L_q will usually be one of the first values you will want to determine in problem solving.

EXAMPLE 1

Customers arrive at a bakery at an average rate of 18 per hour on weekday mornings. The arrival distribution can be described by a Poisson distribution with a mean of 16. Each clerk can serve a customer in an average of three minutes; this time can be described by an exponential distribution with a mean of 3.0 minutes.

a. What are the arrival and service *rates*?

b. Compute the average number of customers being served at any time.

c. Suppose it has been determined that the average number of customers waiting in line is 3.2. Compute the average number of customers in the system (i.e., waiting in line or being served), the average time customers wait in line, and the average time in the system.

d. Determine the system utilization for $M = 1, 2,$ and 3 servers.

SOLUTION

a. The arrival rate is given in the problem: $\lambda = 16$ customers per hour. Change the service *time* to a comparable hourly rate by first restating the time in hours and then taking its reciprocal. Thus, (3 minutes per customer)/(60 minutes per hour) $= 1/20 = 1/\mu$. Its reciprocal is $\mu = 20$ customers per hour.

b. $r = \dfrac{\lambda}{\mu} = \dfrac{16}{20} = .80$ customer.

c. Given: $L_q = 3.2$ customers.

$L_s = L_q + r = 3.2 + .80 = 4.0$ customers

$W_q = \dfrac{L_q}{\lambda} = \dfrac{3.2}{16} = .20$ hour, or .20 hour $\times$ 60 minutes/hour $= 12$ minutes

$W_s =$ Waiting in line plus service

$= W_q + \dfrac{1}{\mu} = .20 + \dfrac{1}{20} = .25$ hour, or 15 minutes

d. System utilization is $\rho = \dfrac{\lambda}{M\mu}$.

For $M = 1, \rho = \dfrac{16}{1(20)} = .80$

For $M = 2, \rho = \dfrac{16}{2(20)} = .40$

For $M = 3, \rho = \dfrac{16}{3(20)} = .27$

Note that as the system capacity as measured by $M\mu$ increases, the system utilization for a given arrival rate decreases.

Single Server, Exponential Service Time, M/M/1[1]

The simplest model involves a system that has one server (or a single crew). The queue discipline is first-come, first-served, and it is assumed that the customer arrival rate can be approximated by a Poisson distribution and service time by a negative exponential distribution. There is no limit on length of queue.

Table 18.2 lists the formulas for the single-channel model, which should be used in conjunction with formulas 18–1 through 18–5.

---

[1]This notation is commonly used to specify waiting line models. The first symbol refers to arrivals, the second to service, and the third to the number of servers. M stands for a rate that can be described by a Poisson distribution, or, equivalently, a time that can be described by an exponential distribution. Hence, M/M/1 indicates a Poisson arrival rate, a Poisson service rate, and one server. The symbol D is used to denote a deterministic (i.e., constant) service rate. Thus the notation M/D/1 would indicate the arrival rate is Poisson and the service rate is constant. Finally, the notation M/M/s would indicate multiple servers.

TABLE 18.2
Formulas for basic single-server model

Performance Measure	Equation	
Average number in line	$L_q = \dfrac{\lambda^2}{\mu(\mu - \lambda)}$	(18–6)
Probability of zero units in the system	$P_0 = 1 - \left(\dfrac{\lambda}{\mu}\right)$	(18–7)
Probability of n units in the system	$P_n = P_0 \left(\dfrac{\lambda}{\mu}\right)^n$	(18–8a)
Probability of less than n units in the system	$P_{<n} = 1 - \left(\dfrac{\lambda}{\mu}\right)^n$	(18–8b)

EXAMPLE 2

An airline is planning to open a satellite ticket desk in a new shopping plaza, staffed by one ticket agent. It is estimated that requests for tickets and information will average 15 per hour, and requests will have a Poisson distribution. Service time is assumed to be exponentially distributed. Previous experience with similar satellite operations suggests that mean service time should average about three minutes per request. Determine each of the following:

a. System utilization.

b. Percentage of time the server (agent) will be idle.

c. The expected number of customers waiting to be served.

d. The average time customers will spend in the system. .

e. The probability of zero customers in the system and the probability of four customers in the system.

SOLUTION ————

$\lambda = 15$ customers per hour

$$\mu = \frac{1}{\text{Service time}} = \frac{1 \text{ customer}}{3 \text{ minutes}} \times 60 \text{ minutes per hour}$$

$$= 20 \text{ customers per hour}$$

a. $\rho = \dfrac{\lambda}{M\mu} = \dfrac{15}{1(20)} = .75$

b. Percentage idle time $= 1 - \rho = 1 - .75 = .25$, or 25 percent

c. $L_q = \dfrac{\lambda^2}{\mu(\mu - \lambda)} = \dfrac{15^2}{20(20 - 15)} = 2.25$ customers

d. $W_s = \dfrac{L_q}{\lambda} + \dfrac{1}{\mu} = \dfrac{2.25}{15} + \dfrac{1}{20} = .20$ hour, or 12 minutes

e. $P_0 = 1 - \dfrac{\lambda}{\mu} = 1 - \dfrac{15}{20} = .25$ and $P_4 = P_0 \left(\dfrac{\lambda}{\mu}\right)^4 = .25 \left(\dfrac{15}{20}\right)^4 = .079$

Single Server, Constant Service Time, M/D/1

As noted previously, waiting lines are a consequence of random, highly variable arrival and service rates. If a system can reduce or eliminate the variability of either or both, it can shorten waiting lines noticeably. A case in point is a system with constant service time. The effect of a constant service time is to cut in half the average number of customers waiting in line:

$$L_q = \frac{\lambda^2}{2\mu(\mu - \lambda)} \tag{18–9}$$

The average time customers spend waiting in line is also cut in half. Similar improvements can be realized by smoothing arrival times (e.g., by use of appointments).

EXAMPLE 3   Wanda's Car Wash & Dry is an automatic, five-minute operation with a single bay. On a typical Saturday morning, cars arrive at a mean rate of eight per hour, with arrivals tending to follow a Poisson distribution. Find

a.   The average number of cars in line.

b.   The average time cars spend in line and service.

SOLUTION

λ = 8 cars per hour

μ = 1 per 5 minutes, or 12 per hour

a. $L_q = \dfrac{\lambda^2}{2\mu(\mu - \lambda)} = \dfrac{8^2}{2(12)(12 - 8)} = .667$ car

b. $W_s = \dfrac{L_q}{\lambda} + \dfrac{1}{\mu} = \dfrac{.667}{8} + \dfrac{1}{12} = .167$ hour, or 10 minutes

Multiple Servers, M/M/S

A multiple-channel system exists whenever there are two or more servers working *independently* to provide service to customer arrivals. Use of the model involves the following assumptions:

1. A Poisson arrival rate and exponential service time.
2. Servers all work at the same average rate.
3. Customers form a single waiting line (in order to maintain first-come, first-served processing).

Formulas for the multiple-channel model are listed in Table 18.3. Obviously, the multiple-channel formulas are more complex than the single-channel formulas, especially the formulas for L_q and P_0. These formulas are shown primarily for completeness; you can actually determine their values using Table 18.4, which gives values of L_q and P_0 for selected values of λ/μ and M.

To use Table 18.4, compute the value of λ/μ and round according to the number of decimal places given for that ratio in the table. Then simply read the values of L_q and P_0 for the appropriate number of channels, M. For instance, if $\lambda/\mu = 0.50$ and $M = 2$, the table provides a value of 0.033 for L_q and a value of .600 for P_0. These values can then be used to compute other measures of system performance. Note that the formulas in Table 18.3 and the values in

TABLE 18.3
Multiple-server queuing formulas

Performance Measure	Equation	
Average number in line	$L_q = \dfrac{\lambda\mu\left(\dfrac{\lambda}{\mu}\right)^M}{(M-1)!(M\mu - \lambda)^2}P_0$	(18–10)
Probability of zero units in the system	$P_0 = \left[\sum_{n=0}^{M-1}\dfrac{\left(\dfrac{\lambda}{\mu}\right)^n}{n!} + \dfrac{\left(\dfrac{\lambda}{\mu}\right)^M}{M!\left(1 - \dfrac{\lambda}{M\mu}\right)}\right]^{-1}$	(18–11)
Average waiting time for an arrival not immediately served	$W_a = \dfrac{1}{M\mu - \lambda}$	(18–12)
Probability that an arrival will have to wait for service	$P_w = \dfrac{W_q}{W_a}$	(18–13)

TABLE 18.4
Infinite-source values for L_q and P_0 given λ/μ and M

λ/μ	M	L_q	P_0	λ/μ	M	L_q	P_0	λ/μ	M	L_q	P_0
0.15	1	0.026	.850	1.1	2	0.477	.290	2.4	3	2.589	.056
	2	0.001	.860		3	0.066	.327		4	0.431	.083
0.20	1	0.050	.800		4	0.011	.332		5	0.105	.089
	2	0.002	.818	1.2	2	0.675	.250		6	0.027	.090
0.25	1	0.083	.750		3	0.094	.294		7	0.007	.091
	2	0.004	.778		4	0.016	.300	2.5	3	3.511	.045
0.30	1	0.129	.700		5	0.003	.301		4	0.533	.074
	2	0.007	.739	1.3	2	0.951	.212		5	0.130	.080
0.35	1	0.188	.650		3	0.130	.264		6	0.034	.082
	2	0.011	.702		4	0.023	.271		7	0.009	.082
0.40	1	0.267	.600		5	0.004	.272	2.6	3	4.933	.035
	2	0.017	.667	1.4	2	1.345	.176		4	0.658	.065
0.45	1	0.368	.550		3	0.177	.236		5	0.161	.072
	2	0.024	.633		4	0.032	.245		6	0.043	.074
	3	0.002	.637		5	0.006	.246		7	0.011	.074
0.50	1	0.500	.500	1.5	2	1.929	.143	2.7	3	7.354	.025
	2	0.033	.600		3	0.237	.211		4	0.811	.057
	3	0.003	.606		4	0.045	.221		5	0.198	.065
0.55	1	0.672	.450		5	0.009	.223		6	0.053	.067
	2	0.045	.569	1.6	2	2.844	.111		7	0.014	.067
	3	0.004	.576		3	0.313	.187	2.8	3	12.273	.016
0.60	1	0.900	.400		4	0.060	.199		4	1.000	.050
	2	0.059	.538		5	0.012	.201		5	0.241	.058
	3	0.006	.548	1.7	2	4.426	.081		6	0.066	.060
0.65	1	1.207	.350		3	0.409	.166		7	0.018	.061
	2	0.077	.509		4	0.080	.180	2.9	3	27.193	.008
	3	0.008	.521		5	0.017	.182		4	1.234	.044
0.70	1	1.633	.300	1.8	2	7.674	.053		5	0.293	.052
	2	0.098	.481		3	0.532	.146		6	0.081	.054
	3	0.011	.495		4	0.105	.162		7	0.023	.055
0.75	1	2.250	.250		5	0.023	.165	3.0	4	1.528	.038
	2	0.123	.455	1.9	2	17.587	.026		5	0.354	.047
	3	0.015	.471		3	0.888	.128		6	0.099	.049
0.80	1	3.200	.200		4	0.136	.145		7	0.028	.050
	2	0.152	.429		5	0.030	.149		8	0.008	.050
	3	0.019	.447		6	0.007	.149	3.1	4	1.902	.032
0.85	1	4.817	.150	2.0	3	0.889	.111		5	0.427	.042
	2	0.187	.404		4	0.174	.130		6	0.120	.044
	3	0.024	.425		5	0.040	.134		7	0.035	.045
	4	0.003	.427		6	0.009	.135		8	0.010	.045
0.90	1	8.100	.100	2.1	3	1.149	.098	3.2	4	2.386	.027
	2	0.229	.379		4	0.220	.117		5	0.513	.037
	3	0.030	.403		5	0.052	.121		6	0.145	.040
	4	0.004	.406		6	0.012	.122		7	0.043	.040
0.95	1	18.050	.050	2.2	3	1.491	.081		8	0.012	.041
	2	0.277	.356		4	0.277	.105	3.3	4	3.027	.023
	3	0.037	.383		5	0.066	.109		5	0.615	.033
	4	0.005	.386		6	0.016	.111		6	0.174	.036
1.0	2	0.333	.333	2.3	3	1.951	.068		7	0.052	.037
	3	0.045	.364		4	0.346	.093		8	0.015	.037
	4	0.007	.367		5	0.084	.099				
					6	0.021	.100				

(continued)

TABLE 18.4
(Concluded)

λ/μ	M	L_q	P_0	λ/μ	M	L_q	P_0	λ/μ	M	L_q	P_0
3.4	4	3.906	.019	4.3	7	0.289	.130	5.2	6	4.301	.003
	5	0.737	.029		8	0.097	.013		7	1.081	.005
	6	0.209	.032		9	0.033	.014		8	0.368	.005
	7	0.063	.033		10	0.011	.014		9	0.135	.005
	8	0.019	.033	4.4	5	5.268	.006		10	0.049	.005
3.5	4	5.165	.015		6	1.078	.010		11	0.018	.006
	5	0.882	.026		7	0.337	.012	5.3	6	5.303	.003
	6	0.248	.029		8	0.114	.012		7	1.249	.004
	7	0.076	.030		9	0.039	.012		8	0.422	.005
	8	0.023	.030		10	0.013	.012		9	0.155	.005
	9	0.007	.030	4.5	5	6.862	.005		10	0.057	.005
3.6	4	7.090	.011		6	1.265	.009		11	0.021	.005
	5	1.055	.023		7	0.391	.010		12	0.007	.005
	6	0.295	.026		8	0.134	.011	5.4	6	6.661	.002
	7	0.019	.027		9	0.046	.011		7	1.444	.004
	8	0.028	.027		10	0.015	.011		8	0.483	.004
	9	0.008	.027	4.6	5	9.289	.004		9	0.178	.004
3.7	4	10.347	.008		6	1.487	.008		10	0.066	.004
	5	1.265	.020		7	0.453	.009		11	0.024	.005
	6	0.349	.023		8	0.156	.010		12	0.009	.005
	7	0.109	.024		9	0.054	.010	5.5	6	8.590	.002
	8	0.034	.025		10	0.018	.010		7	1.674	.003
	9	0.010	.025	4.7	5	13.382	.003		8	0.553	.004
3.8	4	16.937	.005		6	1.752	.007		9	0.204	.004
	5	1.519	.017		7	0.525	.008		10	0.077	.004
	6	0.412	.021		8	0.181	.009		11	0.028	.004
	7	0.129	.022		9	0.064	.009		12	0.010	.004
	8	0.041	.022		10	0.022	.009	5.6	6	11.519	.001
	9	0.013	.022	4.8	5	21.641	.002		7	1.944	.003
3.9	4	36.859	.002		6	2.071	.006		8	0.631	.003
	5	1.830	.015		7	0.607	.008		9	0.233	.004
	6	0.485	.019		8	0.209	.008		10	0.088	.004
	7	0.153	.020		9	0.074	.008		11	0.033	.004
	8	0.050	.020		10	0.026	.008		12	0.012	.004
	9	0.016	.020	4.9	5	46.566	.001	5.7	6	16.446	.001
4.0	5	2.216	.013		6	2.459	.005		7	2.264	.002
	6	0.570	.017		7	0.702	.007		8	0.721	.003
	7	0.180	.018		8	0.242	.007		9	0.266	.003
	8	0.059	.018		9	0.087	.007		10	0.102	.003
	9	0.019	.018		10	0.031	.007		11	0.038	.003
4.1	5	2.703	.011		11	0.011	.007		12	0.014	.003
	6	0.668	.015	5.0	6	2.938	.005	5.8	6	26.373	.001
	7	0.212	.016		7	0.810	.006		7	2.648	.002
	8	0.070	.016		8	0.279	.006		8	0.823	.003
	9	0.023	.017		9	0.101	.007		9	0.303	.003
4.2	5	3.327	.009		10	0.036	.007		10	0.116	.003
	6	0.784	.013		11	0.013	.007		11	0.044	.003
	7	0.248	.014	5.1	6	3.536	.004		12	0.017	.003
	8	0.083	.015		7	0.936	.005	5.9	6	56.300	.000
	9	0.027	.015		8	0.321	.006		7	3.113	.002
	10	0.009	.015		9	0.117	.006		8	0.939	.002
4.3	5	4.149	.008		10	0.042	.006		9	0.345	.003
	6	0.919	.012		11	0.015	.006		10	0.133	.003

Table 18.4 yield *average* amounts (i.e., expected values). Note also that Table 18.4 can also be used for some single-channel problems (i.e., $M = 1$).

Alpha Taxi and Hauling Company has seven cabs stationed at the airport. The company has determined that during the late-evening hours on weeknights, customers request cabs at a rate that follows the Poisson distribution with a mean of 6.6 per hour. Service time is exponential with a mean of 50 minutes per customer. Assume that there is one customer per cab. Find each of the performance measures listed in Table 18.3 and the system utilization.

EXAMPLE 4

SOLUTION ——————

$$\lambda = 6.6 \text{ per hour} \qquad M = 7 \text{ cabs (servers)}$$

$$\mu = \frac{1 \text{ customer per trip}}{50 \text{ minutes per trip} \div 60 \text{ minutes per hour}}$$

$$= 1.2 \text{ customers per hour per cab}$$

$\lambda/\mu = 5.5$. From Table 18.4 with $M = 7$, $L_q = 1.674$ and $P_0 = .003$

a. $L_q = 1.674$ customers

b. $P_0 = .003$

c. $W_a = \dfrac{1}{M\mu - \lambda} = \dfrac{1}{7(1.2) - 6.6} = .556 \text{ hour, or } 33.36 \text{ minutes}$

d. $W_q = \dfrac{L_q}{\lambda} = \dfrac{1.674}{6.6} = .2536 \text{ hour, or } 15.22 \text{ minutes}$

$P_W = \dfrac{W_q}{W_a} = \dfrac{.2536}{.556} = .456$

e. $\rho = \dfrac{\lambda}{M\mu} = \dfrac{6.6}{7(1.2)} = .786$

The Excel template also can be used to solve Example 4. After entering $\lambda = 6.6$ and $\mu = 1.2$ at the top of the template, the queuing statistics for 7 servers are shown in the first column of the table in the template. The template also provides queuing statistics for 8 through 12 servers for comparison, although these are not required for this example. In addition, the template can be used to increment λ, μ, or the number of servers to further investigate the queuing system.

Excel

Multiple Channel Waiting Line Model							
Arrival rate	$\lambda =$	6.6			Service rate	$\mu =$	1.2
Increment	$\Delta\lambda =$	0.1			Increment	$\Delta\mu =$	0.1
Interarrival Time	$1/\lambda =$	0.1515			Service time	$1/\mu =$	0.8333
Number of servers (max 12)	$M =$	7	8	9	10	11	12
System Utilization	$\rho =$	0.7857	0.6875	0.6111	0.5500	0.5000	0.4583
Probability system is empty	$P_0 =$	0.0032	0.0038	0.0040	0.0040	0.0041	0.0041
Probability arrival must wait	$P_w =$	0.4564	0.2512	0.1298	0.0628	0.0284	0.0121
Average number in line	$L_q =$	1.6736	0.5527	0.2039	0.0767	0.0284	0.0102
Average number in system	$L_s =$	7.1736	6.0527	5.7039	5.5767	5.5284	5.5102
Average time in line	$W_q =$	0.2536	0.0837	0.0309	0.0116	0.0043	0.0015
Average time in system	$W_s =$	1.0869	0.9171	0.8642	0.8450	0.8376	0.8349
Average waiting time	$W_a =$	0.5556	0.3333	0.2381	0.1852	0.1515	0.1282

The process also can be worked in reverse; that is, an analyst can determine the capacity needed to achieve specified levels of various performance measures. This approach is illustrated in the following example.

EXAMPLE 5

Alpha Taxi and Hauling also plans to have cabs at a new rail station. The expected arrival rate is 4.8 customers per hour, and the service rate (including return time to the rail station) is expected to be 1.5 per hour. How many cabs will be needed to achieve an average time in line of 20 minutes or less?

SOLUTION

λ = 4.8 customers per hour

μ = 1.5 customers per hour

M = ?

$r = \dfrac{\lambda}{\mu} = \dfrac{4.8}{1.5} = 3.2$ customers

W_q (desired) = 20 minutes, or .333 hour

Using $L_q = \lambda \times W_q$, you can solve for L_q: 4.8/hour (.333 hour) = 1.6 units. Thus, the average number waiting should not exceed 1.6 customers. Referring to Table 18.4, with r = 3.2, L_q = 2.386 for M = 4 and 0.513 for M = 5. Hence, five cabs will be needed.

Finally, note that in a situation where there are multiple servers, each with a separate line (e.g., a supermarket), each line would be treated as a single-server system.

Cost Analysis

The design of a service system often reflects the desire of management to balance the cost of capacity with the expected cost of customers waiting in the system. (Note that customer waiting cost refers to the costs incurred by the organization due to customer waiting.) For example, in designing loading docks for a warehouse, the cost of docks plus loading crews must be balanced against the cost of trucks and drivers that will be in the system, both while waiting to be unloaded and while actually being unloaded. Similarly, the cost of having a mechanic wait for tools at a tool crib must be balanced against the cost of servers at the crib. In cases where the customers are not employees (e.g., retail sales), the costs can include lost sales when customers refuse to wait, the cost of providing waiting space, and the cost of added congestion (lost business, shoplifting).

The optimal capacity (usually in terms of number of channels) is one that minimizes the sum of customer waiting costs and capacity or server costs. Thus, the goal is

Minimize $\dfrac{\text{Total}}{\text{cost}} = \dfrac{\text{Customer}}{\text{waiting cost}} + \dfrac{\text{Capacity}}{\text{cost}}$

The simplest approach to a cost analysis involves computing *system* costs, that is, computing the costs for customers in the system and total capacity cost.

An iterative process is used to identify the capacity size that will minimize total costs. Capacity is incremented one unit at a time (e.g., increase the number of channels by one) and the total cost is computed at each increment. Because the total cost curve is U-shaped, usually the total cost will initially decrease as capacity is increased and then it will eventually begin to increase. Once it begins to increase, additional increases in capacity will cause it to continue to increase. Hence, once that occurs, the optimal capacity size can be readily identified.

The computation of customer waiting costs is based on the average *number* of customers in the *system*. This is perhaps not intuitively obvious; instead, it might seem that customer waiting *time* in the system would be more appropriate. However, that approach would pertain to only *one* customer—it would not convey information concerning *how many* customers would wait that long. Obviously, an average of five customers waiting would involve a lower waiting cost than an average of nine. Therefore, it is necessary to focus on the *number* waiting. Moreover, if an average of two customers are in the system, this is equivalent to having *exactly* two

customers in the system at all times, even though in reality there will be times when zero, one, two, three, or more customers are in the system.

EXAMPLE 6

Trucks arrive at a warehouse at a rate of 15 per hour during business hours. Crews can unload the trucks at a rate of five per hour. The high unloading rate is due to cargo being containerized. Recent changes in wage rates have caused the warehouse manager to reexamine the question of how many crews to use. The new rates are: crew and dock cost is $100 per hour; truck and driver cost is $120 per hour.

SOLUTION

(L_q values are from Table 18.4 using $\dfrac{\lambda}{\mu} = \dfrac{15}{5} = 3.0$.)

Number of Crews	Crew and Dock Cost	$\left[L_s = L_q + \dfrac{\lambda}{\mu}\right]$ Average Number in System	[$L_s \times$ \$120] Driver/Truck Cost	Total Cost
4	\$400	1.528 + 3.0 = 4.528	\$543.36	\$943.36
5	500	.354 + 3.0 = 3.354	402.48	902.48 (minimum)
6	600	.099 + 3.0 = 3.099	371.88	971.88
7	700	.028 + 3.0 = 3.028	363.36	1,063.36

Five crews will minimize the total cost. Because the total cost will continue to increase once the minimum is reached, it is not really necessary to compute total costs for crew sizes larger than six, because total cost increased as the crew size was increased from five to six, indicating that a crew of five is optimal.

One additional point should be made concerning cost analysis. Because both customer waiting costs and capacity costs often reflect estimated amounts, the apparent optimal solution may not represent the true optimum. One ramification of this is that when computations are shown to the nearest penny, or even the nearest dollar, the total cost figures may seem to imply a higher degree of precision than is really justified by the cost estimates. This is compounded by the fact that arrival and service rates may either be approximations or not be exactly represented by the Poisson/exponential distributions. Another ramification is that if cost estimates can be obtained as *ranges* (e.g., customer waiting cost is estimated to range between $40 and $50 per hour), total costs should be computed using both ends of the range to see whether the optimal solution is affected. If it is, management must decide whether to expend additional effort to obtain more precise cost estimates or choose one of the two indicated optimal solutions. Management would most likely choose to employ the latter strategy if there were little disparity between total costs of various capacity levels close to the indicated optimal solutions.

Maximum Line Length

Another question that often comes up in capacity planning is the amount of space to allocate for waiting lines. Theoretically, with an infinite population source, the waiting line can become infinitely long. This implies that no matter how much space is allocated for a waiting line, one can never be completely sure that the space requirements won't exceed that amount. Nonetheless, as a practical matter, one can determine a line length that will not be exceeded a specified proportion of the time. For instance, an analyst may wish to know the length of line that will probably not be exceeded 98 percent of the time, or perhaps 99 percent of the time, and use that number as a planning value.

The approximate line length that will satisfy a specified percentage can be determined by solving the following equation for n:

$$n = \frac{\log K}{\log \rho} \quad \text{or} \quad \frac{\ln K}{\ln \rho} \quad \text{where } K = \frac{1 - \dfrac{\text{Specified}}{\text{percentage}}}{L_q(1 - \rho)} \quad (18\text{-}14)$$

The resulting value of n will not usually be an integer. Generally, round *up* to the next integer and treat the value as n. However, as a practical matter, if the computed value of n is less than 0.10 above the next lower integer, round down. Thus, 15.2 would be rounded to 16, but 15.06 would be rounded to 15.

EXAMPLE 7

Determine the maximum length of a waiting line for specified probabilities of 95 percent and 98 percent, for a system in which $M = 2$, $\lambda = 8$ per hour, and $\mu = 5$ per hour.

SOLUTION

$$r = \frac{8}{5} = 1.6 \quad \text{and} \quad \rho = \frac{8}{2(5)} = .80$$

From Table 18.4, $L_q = 2.844$ customers. For 95 percent, using Formula 18–14:

$$K = \frac{1 - .95}{2.844(1 - .80)} = .088$$

$$n = \frac{\ln .088}{\ln .80} = \frac{-2.4304}{-.2231} = 10.89, \text{ which rounds to } 11$$

For 98 percent:

$$K = \frac{1 - .98}{2.844(1 - .80)} = .035$$

$$n = \frac{\ln .035}{\ln .80} = \frac{-3.352}{-.2231} = 15.02, \text{ which rounds to } 15$$

Multiple Priorities

In many queuing systems, processing occurs on a first-come, first-served basis. However, there are situations in which that rule is inappropriate. The reason is that the waiting cost or penalty incurred is not the same for all customers. In a hospital emergency waiting room, a wide variety of injuries and illnesses need treatment. Some may be minor (e.g., sliver in finger) and others may be much more serious, even life-threatening. It is more reasonable to treat the most serious cases first, letting the nonserious cases wait until all serious cases have been treated. Similarly, computer processing of jobs often follows rules other than first-come, first-served (e.g., shortest job first). In such cases, a **multiple-priority** model is useful for describing customer waiting times.

multiple-priority model
Customers are processed according to some measure of importance.

In these systems, arriving customers are assigned to one of several *priority classes,* or categories, according to a predetermined assignment method (e.g., in a hospital emergency room, heart attacks, serious injuries, and unconscious persons are assigned to the highest priority class; sprains, minor cuts, bruises, and rashes are assigned to the lowest class; and other problems are assigned to one or more intermediate classes). Customers are then processed by class, highest class first. Within each class, processing is first-come, first-served. Thus, all customers in the highest class would be processed before those in the next lower class, then processing would move to that class, and then to the next lower class. Exceptions would occur only if a higher-priority customer arrived; that customer would be processed *after* the customer currently being processed (i.e., service would not be *preemptive*).

This model incorporates all of the assumptions of the basic multiple-server model except that it uses priority serving instead of first-come, first-served. Arrivals to the system are assigned a priority as they arrive (e.g., highest priority = 1, next priority class = 2, next priority class = 3, and so on). An existing queue might look something like this:

TABLE 18.5
Multiple-channel priority service model
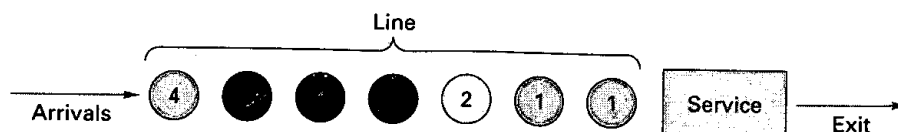
Performance Measure	Formula	Formula Number
System utilization	$\rho = \dfrac{\lambda}{M\mu}$	(18–15)
Intermediate values (L_q from Table 18.4)	$A = \dfrac{\lambda}{(1-\rho)L_q}$	(18–16)
	$B_k = 1 - \sum_{c=1}^{k} \dfrac{\lambda_c}{M\mu}$	(18–17)
	$(B_0 = 1)$	
Average waiting time in line for units in kth priority class	$W_k' = \dfrac{1}{A \cdot B_{k-1} \cdot B_k}$	(18–18)
Average time in the system for units in the kth priority class	$W = W_k' + \dfrac{1}{\mu}$	(18–19)
Average number waiting in line for units in the kth priority class	$L_k = \lambda_k \times W_k'$	(18–20)

Within each class, waiting units are processed in the order they arrived (i.e., first-come, first-served). Thus, in this sequence, the first 1 would be processed as soon as a server was available. The second 1 would be processed when that server or another one became available. If, in the interim, another 1 arrived, it would be next in line *ahead of the first 2*. If there were no new arrivals, the only 2 would be processed by the next available server. At that point, if a new 1 or 2 arrived, it would be processed ahead of the 3s and the 4. Conversely, if a new 4 arrived, it would take its place at the end of the line.

Obviously, a unit with a low priority could conceivably wait a rather long time for processing. In some cases, units that have waited more than some specified time are reassigned to a higher priority.

Table 18.5 gives the appropriate formulas for this multiple-channel priority service model. However, due to the extent of computations involved, it is best to use the appropriate Excel template on the CD for computations.

EXAMPLE 8

A machine shop handles tool repairs in a large company. As each job arrives in the shop, it is assigned a priority based on urgency of the need for that tool. Requests for repair can be described by a Poisson distribution. Arrival rates are: $\lambda_1 = 2$ per hour, $\lambda_2 = 2$ per hour, and $\lambda_3 = 1$ per hour. The service rate is one tool per hour for each server, and there are six servers in the shop. Determine the following information.

a. The system utilization.
b. The average time a tool in each of the priority classes will wait for service.
c. The average time a tool spends in the system for each priority class.
d. The average number of tools waiting for repair in each class.

SOLUTION

$\lambda = \Sigma\lambda_k = 2 + 2 + 1 = 5$ per hour

$M = 6$ servers

$\mu = 1$ customer per hour

Using the Excel template, the solution would appear as follows.
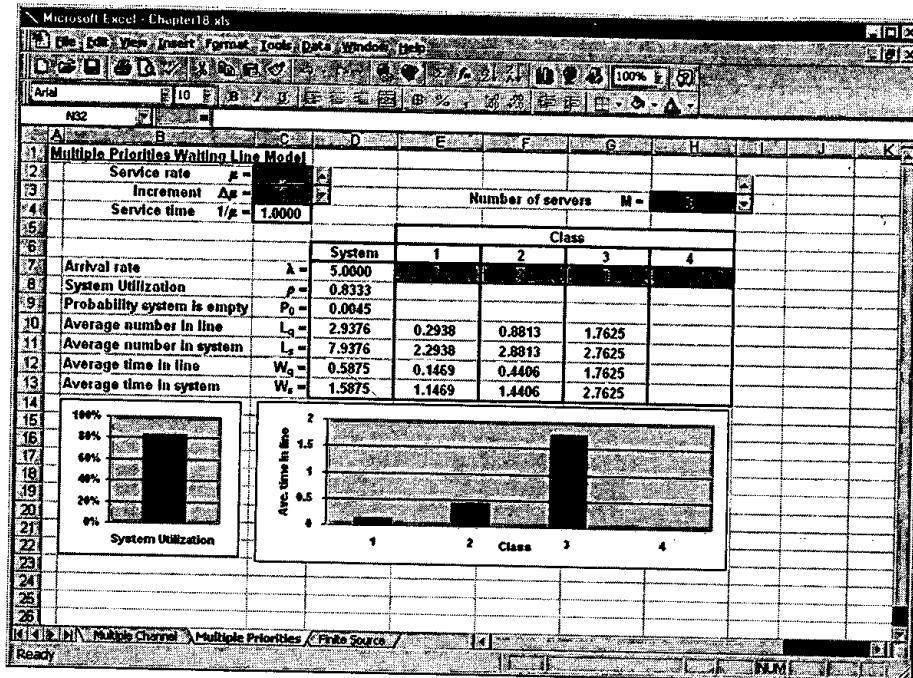
Revising Priorities If any of the waiting times computed in Example 8 are deemed too long by management (e.g., a waiting time of .147 hour for tools in the first class might be too long), there are several options. One is to increase the number of servers. Another is to attempt to increase the service rate, say, by introducing new methods. If such options are not feasible, another approach is to reexamine the membership of each of the priority classifications because, if some repair requests in the first priority class, for example, can be reassigned to the second priority class, this will tend to decrease the average waiting times for repair jobs in the highest priority classification, simply because the arrival rate of those items will be lower.

EXAMPLE 9

The manager of the repair shop, after consulting with the managers of the departments that use the shop's services, has revised the list of tools that are given the highest priorities. This is reflected by revised arrival rates. Suppose that the revised rates are: $\lambda_1 = 1.5$, $\lambda_2 = 2.5$, and λ_3 remains unchanged at 1.0. Determine the following information:

a. The system utilization.

b. The average waiting time for units in each priority class.

SOLUTION

$$\lambda = \Sigma\lambda_k = 1.5 + 2.5 + 1.0 = 5.0$$
$$M = 6$$
$$\mu = 1$$

Example 9 offers several interesting results. One is that through reduction of the arrival rate of the highest priority class, the average waiting time for units in that class has decreased. In other words, removing some members of the highest class and placing them into the next lower class reduced the average waiting time for units that remained in the highest class. Note that the average waiting time for the second priority class also was reduced, even though units were added to that class. Although this may appear counterintuitive, it is necessary to recognize that the *total* waiting time (when all arrivals are taken into account) will remain unchanged. We can see this by noticing that the average *number* waiting (see Example 8, part *d*) is $.294 + .884 + 1.765 = 2.943$. In Example 9, using the average waiting times just computed, the average number waiting in all three classes is

$$\sum_{k=1}^{3} \lambda_k W_k = 1.5(.131) + 2.5(.393) + 1.0(1.765) = 2.944$$

Aside from a slight difference due to rounding, the totals are the same.

Another interesting observation is that the average waiting time for customers in the third priority class did not change from the preceding example. The reason for this is that the *total* arrival rate for the two higher-priority classes did not change, and the average arrival rate for this class did not change. Hence, units assigned to the lowest class must still contend with a combined arrival rate of 4 for the two higher-priority classes.

QUEUING MODEL: FINITE-SOURCE

The finite-source model is appropriate for cases in which the calling population is limited to a relatively small number of potential calls. For instance, one person may be responsible for handling breakdowns on 15 machines; thus, the size of the calling population is 15. However, there may be more than one server or channel; for example, due to a backlog of machines awaiting repairs, the manager might authorize an additional person to work on repairs.

As in the infinite-source models, arrival rates are required to be Poisson and service times exponential. A major difference between the finite- and infinite-source models is that the arrival rate of customers in a finite situation is *affected by* the length of the waiting line; the arrival rate decreases as the length of the line increases simply because there is a decreasing proportion of the population left to generate calls for service. The limit occurs when *all* of the population are waiting in line; at that point the arrival rate is zero since no additional units can arrive.

Because the mathematics of the finite-source model can be complex, analysts often use finite-queuing tables in conjunction with simple formulas to analyze these systems. Table 18.6 contains a list of the key formulas and definitions. You will find it helpful to study the diagram of a cycle that is presented in the table.

Table 18.7 is an abbreviated finite-queuing table used to obtain values of D and F. (Most of the formulas require a value for F.) In order to use the finite-queuing table, follow this procedure:

1. Identify the values for
    a. N, population size.
    b. M, number of servers/channels.

TABLE 18.6
Finite-source queuing formulas and notation

Performance Measure	Formulas		Notation[†]
Service factor	$X = \dfrac{T}{T + U}$	(18–21)	D = Probability that a customer will have to wait in line
Average number waiting	$L = N(1 - F)$	(18–22)	F = Efficiency factor: 1 − percentage waiting in line
Average waiting time	$W = \dfrac{L(T + U)}{N - L} = \dfrac{T(1 - F)}{XF}$	(18–23)	H = Average number of customers being served
Average number running	$J = NF(1 - X)$	(18–24)	J = Average number of customers not in line or in service
Average number being served	$H = FNX$	(18–25)	L = Average number of customers waiting for service
Number in population	$N = J + L + H$	(18–26)	M = Number of service channels
			N = Number of potential customers
			T = Average service time
			U = Average time between customer service requirements per customer
			W = Average time customers wait in line
			X = Service factor

Cycle

	Not waiting or being served	Waiting	
Average number:	J	L	H
Average time:	U	W	T

$$F = \frac{J + H}{J + L + H} *$$

[*]The purpose of this formula is to provide an understanding of F. Because the value of F is needed to compute J, L, and H, the formulas cannot be used to actually compute F. The finite queuing tables must be used for that purpose.

[†]Adapted from L. G. Peck and R. N. Hazelwood, *Finite Queuing Tables* (New York: John Wiley & Sons, 1958). Reprinted by permission.

    c. T, average service time.
    d. U, average time between calls for service per customer.

2. Compute the service factor, $X = T/(T + U)$.
3. Locate the section of the finite-queuing tables for N.
4. Using the value of X as the point of entry, find the values of D and F that correspond to M.
5. Use the values of N, M, X, D, and F as needed to determine the values of the desired measures of system performance.

EXAMPLE 10

One operator loads and unloads a group of five machines. Service time is exponentially distributed with a mean of 10 minutes per cycle. Machines run for an average of 70 minutes between loading and unloading, and this time is also exponential. Find:

a. The average number of machines waiting for the operator.
b. The expected number of machines running.
c. Average downtime.
d. The probability that a machine will not have to wait for service.

TABLE 18.7 Finite-queuing tables

Population 5

X	M	D	F
.012	1	.060	.999
.019	1	.095	.998
.025	1	.125	.997
.030	1	.149	.996
.034	1	.169	.995
.036	1	.179	.994
.040	1	.199	.993
.042	1	.208	.992
.044	1	.218	.991
.046	1	.228	.990
.050	1	.247	.989
.052	1	.257	.988
.054	1	.266	.987
.056	2	.018	.999
	1	.276	.985
.058	2	.019	.999
	1	.285	.984
.060	2	.020	.999
	1	.295	.983
.062	2	.022	.999
	1	.304	.982
.064	2	.023	.999
	1	.314	.981
.066	2	.024	.999
	1	.323	.979
.068	2	.026	.999
	1	.333	.978
.070	2	.027	.999
	1	.342	.977
.075	2	.031	.999
	1	.365	.973
.080	2	.035	.998
	1	.388	.969
.085	2	.040	.998
	1	.410	.965
.090	2	.044	.998
	1	.432	.960
.095	2	.049	.997
	1	.454	.955
.100	2	.054	.997
	1	.475	.950
.105	2	.059	.997
	1	.496	.945
.110	2	.065	.996
	1	.516	.939
.115	2	.071	.995
	1	.537	.933
.120	2	.076	.995
	1	.556	.927

X	M	D	F
.125	2	.082	.994
	1	.575	.920
.130	2	.089	.993
	1	.594	.914
.135	2	.095	.993
	1	.612	.907
.140	2	.102	.992
	1	.630	.900
.145	3	.011	.999
	2	.109	.991
	1	.647	.892
.150	3	.012	.999
	2	.115	.990
	1	.664	.885
.155	3	.013	.999
	2	.123	.989
	1	.680	.877
.160	3	.015	.999
	2	.130	.988
	1	.695	.869
.165	3	.016	.999
	2	.137	.987
	1	.710	.861
.170	3	.017	.999
	2	.145	.985
	1	.725	.853
.180	3	.021	.999
	2	.161	.983
	1	.752	.836
.190	3	.024	.998
	2	.177	.980
	1	.778	.819
.200	3	.028	.998
	2	.194	.976
	1	.801	.801
.210	3	.032	.998
	2	.211	.973
	1	.822	.783
.220	3	.036	.997
	2	.229	.969
	1	.842	.765
.230	3	.041	.997
	2	.247	.965
	1	.860	.747
.240	3	.046	.996
	2	.265	.960
	1	.876	.730
.250	3	.052	.995
	2	.284	.955
	1	.890	.712

X	M	D	F
.260	3	.058	.994
	2	.303	.950
	1	.903	.695
.270	3	.064	.994
	2	.323	.944
	1	.915	.677
.280	3	.071	.993
	2	.342	.938
	1	.925	.661
.290	4	.007	.999
	3	.079	.992
	2	.362	.932
	1	.934	.644
.300	4	.008	.999
	3	.086	.990
	2	.382	.926
	1	.942	.628
.310	4	.009	.999
	3	.094	.989
	2	.402	.919
	1	.950	.613
.320	4	.010	.999
	3	.103	.988
	2	.422	.912
	1	.956	.597
.330	4	.012	.999
	3	.112	.986
	2	.442	.904
	1	.962	.583
.340	4	.013	.999
	3	.121	.985
	2	.462	.896
	1	.967	.569
.360	4	.017	.998
	3	.141	.981
	2	.501	.880
	1	.975	.542
.380	4	.021	.998
	3	.163	.976
	2	.540	.863
	1	.981	.516
.400	4	.026	.997
	3	.186	.972
	2	.579	.845
	1	.986	.493
.420	4	.031	.997
	3	.211	.966
	2	.616	.826
	1	.989	.471
.440	4	.037	.996

X	M	D	F
.440	3	.238	.960
	2	.652	.807
	1	.992	.451
.460	4	.045	.995
	3	.266	.953
	2	.686	.787
	1	.994	.432
.480	4	.053	.994
	3	.296	.945
	2	.719	.767
	1	.996	.415
.500	4	.063	.992
	3	.327	.936
	2	.750	.748
	1	.997	.399
.520	4	.073	.991
	3	.359	.927
	2	.779	.728
	1	.998	.384
.540	4	.085	.989
	3	.392	.917
	2	.806	.708
	1	.998	.370
.560	4	.098	.986
	3	.426	.906
	2	.831	.689
	1	.999	.357
.580	4	.113	.984
	3	.461	.895
	2	.854	.670
	1	.999	.345
.600	4	.130	.981
	3	.497	.883
	2	.875	.652
	1	.999	.333
.650	4	.179	.972
	3	.588	.850
	2	.918	.608
	1	.999	.308
.700	4	.240	.960
	3	.678	.815
	2	.950	.568
	1	.999	.286
.750	4	.316	.944
	3	.763	.777
	2	.972	.532
.800	4	.410	.924
	3	.841	.739
	2	.987	.500

(continued)

TABLE 18.7 (Continued)

X	M	D	F	X	M	D	F	X	M	D	F	X	M	D	F
.850	4	.522	.900	.064	1	.547	.940	.135	2	.415	.952	.220	5	.030	.998
	3	.907	.702	.066	2	.126	.995		1	.907	.699		4	.124	.990
	2	.995	.470		1	.562	.936	.140	4	.028	.999		3	.366	.954
.900	4	.656	.871	.068	3	.020	.999		3	.132	.991		2	.761	.815
	3	.957	.666		2	.133	.994		2	.437	.947		1	.993	.453
	2	.998	.444		1	.577	.931		1	.919	.680	.230	5	.037	.998
.950	4	.815	.838	.070	3	.022	.999	.145	4	.032	.999		4	.142	.988
	3	.989	.631		2	.140	.994		3	.144	.990		3	.400	.947
Population 10					1	.591	.926		2	.460	.941		2	.791	.794
.016	1	.144	.997	.075	3	.026	.999		1	.929	.662		1	.995	.434
.019	1	.170	.996		2	.158	.992	.150	4	.036	.998	.240	5	.044	.997
.021	1	.188	.995		1	.627	.913		3	.156	.989		4	.162	.986
.023	1	.206	.994	.080	3	.031	.999		2	.483	.935		3	.434	.938
.025	1	.224	.993		2	.177	.990		1	.939	.644		2	.819	.774
.026	1	.232	.992		1	.660	.899	.155	4	.040	.998		1	.996	.416
.028	1	.250	.991	.085	3	.037	.999		3	.169	.987	.250	6	.010	.999
.030	1	.268	.990		2	.196	.988		2	.505	.928		5	.052	.997
.032	2	.033	.999		1	.692	.883		1	.947	.627		4	.183	.983
	1	.285	.988	.090	3	.043	.998	.160	4	.044	.998		3	.469	.929
.034	2	.037	.999		2	.216	.986		3	.182	.986		2	.844	.753
	1	.301	.986		1	.722	.867		2	.528	.921		1	.997	.400
.036	2	.041	.999	.095	3	.049	.998		1	.954	.610	.260	6	.013	.999
	1	.320	.984		2	.237	.984	.165	4	.049	.997		5	.060	.996
.038	2	.046	.999		1	.750	.850		3	.195	.984		4	.205	.980
	1	.337	.982	.100	3	.056	.998		2	.550	.914		3	.503	.919
.040	2	.050	.999		2	.258	.981		1	.961	.594		2	.866	.732
	1	.354	.980		1	.776	.832	.170	4	.054	.997		1	.998	.384
.042	2	.055	.999	.105	3	.064	.997		3	.209	.982	.270	6	.015	.999
	1	.371	.978		2	.279	.978		2	.571	.906		5	.070	.995
.044	2	.060	.998		1	.800	.814		1	.966	.579		4	.228	.976
	1	.388	.975	.110	3	.072	.997	.180	5	.013	.999		3	.537	.908
.046	2	.065	.998		2	.301	.974		4	.066	.996		2	.886	.712
	1	.404	.973		1	.822	.795		3	.238	.978		1	.999	.370
.048	2	.071	.998	.115	3	.081	.996		2	.614	.890	.280	6	.018	.999
	1	.421	.970		2	.324	.971		1	.975	.890		5	.081	.994
.050	2	.076	.998		1	.843	.776	.190	5	.016	.999		4	.252	.972
	1	.437	.967	.120	4	.016	.999		4	.078	.995		3	.571	.896
.052	2	.082	.997		3	.090	.995		3	.269	.973		2	.903	.692
	1	.454	.963		2	.346	.967		2	.654	.873		1	.999	.357
.054	2	.088	.997		1	.861	.756		1	.982	.522	.290	6	.022	.999
	1	.470	.960	.125	4	.019	.999	.200	5	.020	.999		5	.093	.993
.056	2	.094	.997		3	.100	.994		4	.092	.994		4	.278	.968
	1	.486	.956		2	.369	.962		3	.300	.968		3	.603	.884
.058	2	.100	.996		1	.878	.737		2	.692	.854		2	.918	.672
	1	.501	.953	.130	4	.022	.999		1	.987	.497		1	.999	.345
.060	2	.106	.996		3	.110	.994	.210	5	.025	.999	.300	6	.026	.998
	1	.517	.949		2	.392	.958		4	.108	.992		5	.106	.991
.062	2	.113	.996		1	.893	.718		3	.333	.961		4	.304	.963
	1	.532	.945	.135	4	.025	.999		2	.728	.835		3	.635	.872
.064	2	.119	.995		3	.121	.993		1	.990	.474				

(continued)

TABLE 18.7 (Concluded)

X	M	D	F	X	M	D	F	X	M	D	F	X	M	D	F
.300	2	.932	.653	.400	6	.105	.991	.500	3	.972	.598	.650	7	.353	.954
	1	.999	.333		5	.292	.963		2	.999	.400		6	.651	.878
.310	6	.031	.998		4	.591	.887	.520	8	.026	.998		5	.882	.759
	5	.120	.990		3	.875	.728		7	.115	.989		4	.980	.614
	4	.331	.957		2	.991	.499		6	.316	.958		3	.999	.461
	3	.666	.858	.420	7	.034	.993		5	.606	.884	.700	9	.040	.997
	2	.943	.635		6	.130	.987		4	.864	.752		8	.200	.979
.320	6	.036	.998		5	.341	.954		3	.980	.575		7	.484	.929
	5	.135	.988		4	.646	.866		2	.999	.385		6	.772	.836
	4	.359	.952		3	.905	.700	.540	8	.034	.997		5	.940	.711
	3	.695	.845		2	.994	.476		7	.141	.986		4	.992	.571
	2	.952	.617	.440	7	.045	.997		6	.363	.949	.750	9	.075	.994
.330	6	.042	.997		6	.160	.984		5	.658	.867		8	.307	.965
	5	.151	.986		5	.392	.943		4	.893	.729		7	.626	.897
	4	.387	.945		4	.698	.845		3	.986	.555		6	.870	.792
	3	.723	.831		3	.928	.672	.560	8	.044	.996		5	.975	.666
	2	.961	.600		2	.996	.454		7	.171	.982		4	.998	.533
.340	7	.010	.999	.460	8	.011	.999		6	.413	.939	.800	9	.134	.988
	6	.049	.997		7	.058	.995		5	.707	.848		8	.446	.944
	5	.168	.983		6	.193	.979		4	.917	.706		7	.763	.859
	4	.416	.938		5	.445	.930		3	.991	.535		6	.939	.747
	3	.750	.816		4	.747	.822	.580	8	.057	.995		5	.991	.625
	2	.968	.584		3	.947	.646		7	.204	.977		4	.999	.500
.360	7	.014	.999		2	.998	.435		6	.465	.927	.850	9	.232	.979
	6	.064	.995	.480	8	.015	.999		5	.753	.829		8	.611	.916
	5	.205	.978		7	.074	.994		4	.937	.684		7	.879	.818
	4	.474	.923		6	.230	.973		3	.994	.517		6	.978	.705
	3	.798	.787		5	.499	.916	.600	9	.010	.999		5	.998	.588
	2	.978	.553		4	.791	.799		8	.072	.994	.900	9	.387	.963
.380	7	.019	.999		3	.961	.621		7	.242	.972		8	.785	.881
	6	.083	.993		2	.998	.417		6	.518	.915		7	.957	.777
	5	.247	.971	.500	8	.020	.999		5	.795	.809		6	.995	.667
	4	.533	.906		7	.093	.992		4	.953	.663	.950	9	.630	.938
	3	.840	.758		6	.271	.966		3	.996	.500		8	.934	.841
	2	.986	.525		5	.553	.901	.650	9	.021	.999		7	.994	.737
.400	7	.026	.998		4	.830	.775		8	.123	.988				

SOLUTION ———

$N = 5$

$T = 10$ minutes

$M = 1$

$U = 70$ minutes

$$X = \frac{T}{T + U} = \frac{10}{10 + 70} = .125$$

From Table 18.7, with $N = 5$, $M = 1$, and $X = .125$, $D = .575$ and $F = .920$.

a. Average number waiting, $L = N(1 - F) = 5(1 - .920) = .40$ machine.

b. Expected number running, $J = NF(1 - X) = 5(.920)(1 - .125) = 4.025$ machines.

c. Downtime = Waiting time + Service time:

Waiting time, $W = \dfrac{L(T + U)}{N - L} = \dfrac{.40(10 + 70)}{5 - .40} = 6.957$ minutes

Downtime = 6.957 minutes + 10 minutes = 16.957 minutes

d. Probability of not waiting = 1 − Probability of waiting

$$= 1 - D$$

$$= 1 - .575 = .425$$

Using the Excel template, the solution to Example 10 would appear as follows:

Excel

Suppose that in Example 10, operators are paid $10 per hour, and machine downtime costs $16 per hour. Should the department add another operator if the goal is cost optimization?

EXAMPLE 11

Compare the total cost of the present system with the expected total cost of the proposed system:

SOLUTION

M	Average Number Down, N − J	Average Down Cost (per hour), (N − J)$16	Operator Cost (per hour)	Total Cost (per hour)
1	.975	$15.60	$10	$25.60
2	.651	10.42	20	30.42

Hence, the present system is superior because its total cost is less than the expected total cost using two operators.

Hotels Exploring Easier Customer Check-ins   NEWSCLIP

Salina Kahn

Travelers can avoid line by registering in shuttles or at the airport in some places.

The hotel front desk is no longer the only place where travelers check in and get their room keys.

More hotels are shortening front desk lines by doing that outside their front door—on airport shuttle buses and even at airports.

Hotels have long allowed guests to register via the phone or the Internet and then pick up a key at a special counter or kiosk in the lobby. Now hotels are accelerating express check-in by using wireless technology to program electronic card keys outside the hotel.

The improvements should be popular: 81 percent of business travelers say express check-in and checkout services are very desirable, according to the 1999 Business Travel Monitor by Yesawich, Pepperdine & Brown and Yankelovich Partners.

Hilton plans to use the technology to provide curbside check-in at hotels in Honolulu, New York and Anaheim, Calif., early next year, says John Luke, Hilton vice president.

Starwood Hotels and Resorts will test curbside check-in at the Boston Park Plaza within 60 days.

Other hotels where guests check in before they reach the lobby:

- The Fairmont Vancouver Airport Hotel accepts reservations, checks bags and hands out keys from two "satellite lobbies" in the airport.
- Park Place Entertainment plans to open a check-in counter at McCarran International Airport in Las Vegas early next year.
- Hilton Boston Logan Airport began in September to allow guests to register on the shuttle bus ride from the airport to the hotel.
- The Loews Portofino Bay Hotel in Orlando, Fla., provides curbside check-in.

Source: Reprinted with permission of the *Rochester Democrat and Chronicle,* November 7, 1999, p. 1E.

OTHER APPROACHES

The discussion in this chapter has focused on designing service systems that achieve a balance between service capacity and customer waiting time. The implication is that decision makers can determine an appropriate level of service capacity. In certain instances, such an approach may not be practical for a variety of reasons. One is that the system may be currently in operation and indicated design changes may be too costly, or space restrictions may prevent changes from being made reasonably. One alternative that is particularly suited for queuing systems in which the customers are people rather than inanimate objects is to provide some form of diversion so that the waiting time becomes more tolerable. This involves taking into account, and taking steps to reduce, the *perceived waiting time,* as opposed to the *actual waiting time.* For example, magazines and newspapers can be placed in waiting rooms, as is usually the case in doctors' and dentists' offices. Auto repair shops sometimes use radio or television, and airlines may provide in-flight movies to help occupy the time. Airlines also serve meals and snacks, which help to make the time spent waiting (flying) more pleasant. Some other measures include placing mirrors where people wait for elevators, and asking people to fill out forms, which makes waiting somewhat constructive.

To carry this concept one step further, it is sometimes possible to derive some benefit from customer waiting. For instance, supermarkets position impulse items near checkout counters and gain additional sales, banks advertise current rates and place brochures describing bank services within easy reach of waiting customers, and restaurants have bars where customers can relax and spend money while waiting for their tables.

S

The implication in these ideas is that imagination and creativity can often play an important role in system design and that mathematical approaches are not the only ones worth considering.

Operations Strategy

Managers must carefully assess the costs and benefits of various alternatives for capacity of service systems. Working to increase the processing rate may be a worthwhile option instead of increasing the number of servers. New processing equipment and/or processing methods may contribute to this goal. One important factor to consider is the possibility of reducing variability in processing times by increasing the degree of standardization of the service being provided. In fact, managers of all services would be wise to pursue this goal, not only for the benefits of reduced waiting times, but also for benefits of standardizing server training and hence reducing those costs and times, and the potential for increased quality due to the decreased variety in service requirements.

SUMMARY

Analysis of waiting lines can be an important aspect of the design of service systems. Waiting lines have a tendency to form in such systems even though, in a macro sense, the system is underloaded. The arrival of customers at random times and variability of service times combine to create temporary overloads. When this happens, waiting lines appear. By the same token, at other times the servers are idle.

A major consideration in the analysis of queuing systems is whether the number of potential customers is limited (finite source) or whether entry to the system is unrestricted (infinite source). Five basic queuing models are described, four dealing with infinite-source populations and one dealing with finite-source populations. In general, the models assume that customer arrival rates can be described by a Poisson distribution and that service time can be described by a negative exponential distribution.

KEY TERMS

channel, 782
finite source, 782
infinite source, 782

multiple-priority model, 796
queue discipline, 785
queuing theory, 779

SOLVED PROBLEMS

Problem 1

Infinite source. One of the features of a new machine shop will be a well-stocked tool crib. The manager of the shop must decide on the number of attendants needed to staff the crib. Attendants will receive $9 per hour in salary and fringe benefits. Mechanics' time will be worth $30 per hour, which includes salary and fringe benefits plus lost work time caused by waiting for parts. Based on previous experience, the manager estimates requests for parts will average 18 per hour with a service capacity of 20 requests per hour per attendant. How many attendants should be on duty if the manager is willing to assume that arrival and service rates will be Poisson-distributed? (Assume the number of mechanics is very large, so an infinite-source model is appropriate.)

Solution

$\lambda = 18$ per hour

$\mu = 20$ per hour

The solution requires a trial-and-error approach that reveals the total cost of feasible alternatives (i.e., a utilization less than 100 percent) and selection of the lowest-cost alternative. Note that the total-cost curve will always be U-shaped; increase the number of servers until the total cost shows an increase over the previous value. The optimum will be the number of servers that produced the previous total cost value. Thus,

Number of Servers, M	L_q*	$L_q + \dfrac{\lambda}{\mu} = L$	$9M$: Server Cost (per hour)	$30L_s$: Mechanic Cost (per hour)	Total Cost (per hour)
1 .... 8.1		$8.1 + 0.9 = 9.0$	$9	$270	$279
2 .... 0.229		$0.229 + 0.9 = 1.129$	$18	$ 33.87	$ 52†
3 .... 0.03		$0.03 + 0.9 = 0.93$	$27	$ 27.9	$ 55†

*L_q from Table 18.4, with $r = \lambda/\mu = 18/20 = 0.9$.
†Rounded.

Hence, two servers will produce the lowest total cost.
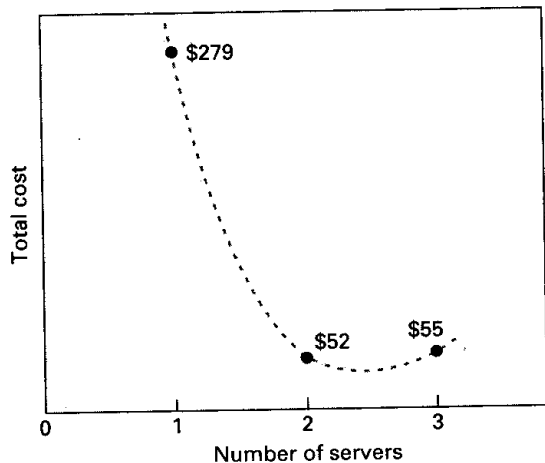
Infinite source. The following is a list of service times for three different operations:

Operation	Service Time
A	8 minutes
B	1.2 hours
C	2 days

a. Determine the service rate for each operation.
b. Would the calculated rates be different if these were interarrival times rather than service times?

Solution

a. The service rate is the reciprocal of service time. Thus, the rates are

A: 1/8 customer per minute = .125 customer per minute, or .125/min. $\times$ 60 min./hr. = 7.5 customers per hour

B: 1/1.2 customer per hour = .833 customer per hour

C: 1/2 customer per day = .50 customer per day

b. No. In either case, the rate is simply the reciprocal of the time.

Finite source. A group of 10 machines is loaded and unloaded by one of three servers. The machines run for an average of six minutes per cycle, and average time to unload and reload is nine minutes. Each time can be described by an exponential distribution. While running, the machines produce at a rate that would yield 16 units per hour if they did not have to pause to wait for a server and be loaded and unloaded. What is the average hourly output of each machine when waiting and serving are taken into account?

Solution

$T = 9$ minutes

$U = 6$ minutes; $x = \dfrac{T}{T+U} = \dfrac{9}{9+6} = .60$

$M = 3$ servers; from Table 18.7, $F = .500$

$N = 10$ machines

a. Compute the average number of machines running:

$J = NF(1 - X) = 10(.500)(.40) = 2$

b. Determine the percentage of machines running, and multiply by output while running:

$$\frac{J}{N} \times (16 \text{ per hour}) = \frac{2}{10} \times (16 \text{ per hour}) = 3.2 \text{ per hour}$$

DISCUSSION AND REVIEW QUESTIONS

1. In what kinds of situations is queuing analysis most appropriate?
2. Why do waiting lines form even though a service system is underloaded?
3. What are the most common measures of system performance in a queuing analysis?
4. What effect would decreasing arrival and service variability have on the effective capacity of a system?
5. What approaches do supermarkets use to offset variations in customer traffic intensity?
6. Contrast *finite* and *infinite* population sources.
7. Will doubling the service rate of a single-channel system reduce the average waiting time in line by a factor of one-half? Explain briefly.
8. In a multiple-channel system, what is the rationale for having customers wait in a single line, as is now being done in many banks and post offices, rather than multiple lines? (*Hint:* The average waiting time is unaffected.)
9. What happens to the length of a waiting line in a highly variable (queuing) setting if a manager attempts to achieve a high percentage of capacity utilization?

TAKING STOCK

1. What general trade-offs are involved in waiting line decisions?
2. Who needs to be involved in assessing the cost of customers waiting for service if the customers are (a) the general public and (b) employees of the organization?
3. How has technology had an impact on analyzing waiting line systems? How has technology improved waiting line performance?

CRITICAL THINKING EXERCISE

The owner of Eat Now Restaurant implemented an expanded menu early last year. The menu was a success, drawing many more customers than the previous menu who seemed to like the increased variety of menu choices. But good news soon became bad news as long waiting lines began to deter customers, and business dropped off. However, because of space and other limitations, there didn't seem to be any viable options to consider. Then a customer mentioned a technique that was being used in the company he worked for, called mass customization. He said it really streamlined processing, and maybe it could work for the restaurant.

Describe how that approach might work at the restaurant and why that could be expected to reduce waiting times. What costs would be involved in transitioning to such a system?

PROBLEMS

1. Repair calls are handled by one repairman at a photocopy shop. Repair time, including travel time, is exponentially distributed, with a mean of two hours per call. Requests for copier repairs come in at a mean rate of three per eight-hour day (assume Poisson). Determine:
   a. The average number of customers awaiting repairs.
   b. System utilization.
   c. The amount of time during an eight-hour day that the repairman is not out on a call.
   d. The probability of two or more customers in the system.

2. A vending machine dispenses hot chocolate or coffee. Service time is 30 seconds per cup and is constant. Customers arrive at a mean rate of 80 per hour, and this rate is Poisson-distributed. Determine:
   a. The average number of customers waiting in line.
   b. The average time customers spend in the system.
   c. The average number in the system.

3. Many of a bank's customers use its automatic teller machine to transact business after normal banking hours. During the early evening hours in the summer months, customers arrive at a certain location at the rate of one every other minute. This can be modeled using a Poisson distribution. Each customer spends an average of 90 seconds completing his or her transactions. Transaction time is exponentially distributed. Determine:

   a. The average time customers spend at the machine, including waiting in line and completing transactions.

   b. The probability that a customer will not have to wait upon arriving at the automatic teller machine.

   c. The average number waiting to use the machine.

4. A small town with one hospital has two ambulances to supply ambulance service. Requests for ambulances during nonholiday weekends average .8 per hour and tend to be Poisson-distributed. Travel and assistance time averages one hour per call and follows an exponential distribution. Find:

   a. System utilization.

   b. The average number of customers waiting.

   c. The average time customers wait for an ambulance.

   d. The probability that *both* ambulances will be busy when a call comes in.

5. The following information pertains to telephone calls to a motel switchboard on a typical Tuesday.

Period	Incoming Rate (calls per minute)	Service Rate (calls per minute per operator)	Number of Operators
Morning.........	1.8	1.5	2
Afternoon........	2.2	1.0	3
Evening	1.4	0.7	3

   a. Determine the average time callers wait to have their calls answered for each period and the probability that a caller will have to wait for each period.

   b. For each case in the previous problem, determine the maximum line length for a probability of 96 percent.

6. Trucks are required to pass through a weighing station so that they can be checked for weight violations. Trucks arrive at the station at the rate of 40 an hour between 7 P.M. and 9 P.M. Currently two inspectors are on duty during those hours, each of whom can inspect 25 trucks an hour.

   a. How many trucks would you expect to see at the weighing station, including those being inspected?

   b. If a truck was just arriving at the station, about how many minutes could the driver expect to be at the station?

   c. What is the probability that both inspectors would be busy at the same time?

   d. How many minutes, on average, would a truck that is not immediately inspected have to wait?

   e. What condition would exist if there was only one inspector?

   f. What is the maximum line length for a probability of .97?

7. The manager of a regional warehouse must decide on the number of loading docks to request for a new facility in order to minimize the sum of dock costs and driver-truck costs. The manager has learned that each driver-truck combination represents a cost of $300 per day and that each dock plus loading crew represents a cost of $1,100 per day.

   a. How many docks should be requested if trucks arrive at the rate of four per day, each dock can handle five trucks per day, and both rates are Poisson?

   b. An employee has proposed adding new equipment that would speed up the loading rate to 5.71 trucks per day. The equipment would cost $100 per day for each dock. Should the manager invest in the new equipment?

8. The parts department of a large automobile dealership has a counter used exclusively for mechanics' requests for parts. The time between requests can be modeled by a negative exponential distribution that has a mean of five minutes. A clerk can handle requests at a rate of 15 per hour, and this can be modeled by a Poisson distribution that has a mean of 15. Suppose there are two clerks at the counter.

   a. On average, how many mechanics would be at the counter, including those being served?

   b. What is the probability that a mechanic would have to wait for service?

   c. If a mechanic has to wait, how long would the average wait be?

   d. What percentage of time are the clerks idle?

   e. If clerks represent a cost of $20 per hour and mechanics a cost of $30 per hour, what number of
      clerks would be optimal in terms of minimizing total cost?

9. One field representative services five customers for a computer manufacturer. Customers request
   assistance at an average (Poisson-distributed) rate of once every four working days. The field repre-
   sentative can handle an average (Poisson-distributed) of one call per day. Determine:
   a. The expected number of customers waiting.
   b. The average length of time customers must wait from the initial request for service until the ser-
      vice has been completed.
   c. The percentage of time the service rep will be idle.
   d. By how much would your answer to part *a* be reduced if a second field rep were added?

10. Two operators handle adjustments for a group of 10 machines. Adjustment time is exponentially dis-
    tributed and has a mean of 14 minutes per machine. The machines operate for an average of 86 min-
    utes between adjustments. While running, each machine can turn out 50 pieces per hour. Find:
    a. The probability that a machine will have to wait for an adjustment.
    b. The average number of machines waiting for adjustment.
    c. The average number of machines being serviced.
    d. The expected hourly output of each machine, taking adjustments into account.
    e. Machine downtime represents a cost of $70 per hour; operator cost (including salary and fringe
       benefits) is $15 per hour. What is the optimum number of operators?

11. One operator services a bank of five machines. Machine running time and service time are both ex-
    ponential. Machines run for an average of 90 minutes between service requirements, and service
    time averages 35 minutes. The operator receives $20 per hour in salary and fringe benefits, and ma-
    chine downtime costs $70 per hour per machine.
    a. If each machine produces 60 pieces per hour while running, find the average hourly output of
       each machine, when waiting and service times are taken into account.
    b. Determine the optimum number of operators.

12. A milling department has 10 machines. Each operates an average of eight hours before requiring ad-
    justment, which takes an average of two hours. While running, each machine can produce 40 pieces
    an hour.
    a. With one adjuster, what is the net average hourly output per machine?
    b. If machine downtime cost is $80 per hour and adjuster cost is $30 per hour, how many adjusters
       would be optimal?

13. Trucks arrive at the loading dock of a wholesale grocer at the rate of 1.2 per hour. A single crew con-
    sisting of two workers can load a truck in about 30 minutes. Crew members receive $10 per hour in
    wages and fringe benefits, and trucks and drivers reflect an hourly cost of $60. The manager is think-
    ing of adding another member to the crew. The service rate would then be 2.4 trucks per hour. As-
    sume rates are Poisson.
    a. Would the third crew member be economical?
    b. Would a fourth member be justifiable if the resulting service capacity were 2.6 trucks per hour?

14. Customers arriving at a service center are assigned to one of three categories, with category 1 given
    the highest priority. Records indicate that an average of nine customers arrive per hour and that one-
    third are assigned to each category. There are two servers, and each can process customers at the rate
    of five per hour. Arrival and service rates can be described by Poisson distributions.
    a. What is the utilization rate for this system?
    b. Determine the average waiting time for units in each class.
    c. Find the average number of customers in each class that are waiting for service.

15. A manager must determine requirements for waiting space for customers. A priority system is used
    to process customers, who are assigned to one of two classes when they enter the processing center.
    The highest-priority class has an arrival rate of four per hour; the other class has an arrival rate of
    two per hour. Both can be described as Poisson-distributed. There are two servers, and each can
    process customers in an average of 15 minutes.
    a. What is the system utilization?
    b. Determine the number of customers of each class that are waiting for service.
    c. Determine the average waiting time for each class.
    d. If the manager could alter the assignment rules so that arrival rates of the two classes were equal,
       what would be the revised average waiting time for each priority class?

16. A priority waiting system assigns arriving customers to one of four classes. Arrival rates (Poisson)
    of the classes are shown in the following table:

Class	Arrivals per Hour
1	2
2	4
3	3
4	2

Five servers process the customers, and each can handle three customers per hour.

a. What is the system utilization?

b. What is the average wait for service by customers in the various classes? How many are waiting in each class, on the average?

c. If the arrival rate of the second priority class could be reduced to three units per hour by shifting some arrivals into the third priority class, how would your answers to part *b* change?

d. What observations can you make based on your answers to part *c*?

17. Referring to Problem 16, suppose that each server could handle four customers per hour. Answer the questions posed in the problem. Explain why the impact of reassigning customers is much less than in Problem 16.

18. During the morning hours at a catalog sales department, telephone calls come in at the rate (Poisson) of 40 per hour. Calls that cannot be answered immediately are put on hold. The system can handle eight callers on hold. If additional calls come in, they receive a busy signal. The three customer service representatives who answer the calls spend an average of three minutes with a customer.

a. What is the probability that a caller will get a busy signal? (*Hint:* Solve for $\log K$ or $\ln K$ using trial and error.)

b. What is the probability that a customer will be put on hold?

Big Bank

CASE

The operations manager of a soon-to-open branch of a large bank is in the process of configuring teller operations. Currently some branches have a separate teller line for customers who have a single transaction, while other branches don't have separate lines. The manager wants to avoid complaints about long waits that have been received at some branches. Because the demographics differ from location to location, a system that works at one branch won't necessarily work at another.

The manager has obtained data on processing times from the bank's home office, and is ready to explore different options for configuring operations. (Fortunately she has her textbook and CD from when she took an operations management course at a nearby university.)

One time that will get special attention is the noon hour on Friday. The plan is to have five tellers available. Under consideration are the following options:

A. Have one waiting line and have the first person in line go to the next available teller.

B. Have two waiting lines: one teller for customers who have a single transaction and four tellers who would handle customers who have multiple transactions.

Processing information: An average of 80 customers are processed during the noon hour. The average processing time for customers with a single transaction is 90 seconds, while the processing time for customers with multiple transactions is 4 minutes. Sixty percent of the customers are expected to have multiple transactions.

If you were the manager, which option would you select? Why? Explain the disparity between the results for the two options. What assumptions did you make in your analysis?

Stopped at a Light? Why Not Read This, You May Have Time

READING

Anna Wilde Mathews

As red signals grow longer, northern Virginia tries an experimental speedup

At the corner of Fairfax County Parkway and Fair Lakes Parkway, drivers see red.

More than four minutes of it.

Hanging over this crowded intersection in northern Virginia's booming suburbs is one of the longest stoplights in the country. Washington-bound commuters curse it. Highflying tech executives making their way to appointments feel grounded. Prepared motorists bring breakfast to nibble while they wait. Occasionally, a gaggle of geese waddles across the road to a nearby pond, slowing things even more.

"You come out here sometimes, and it brings tears to your eyes, it's so bad," says Jeris J. White, the state transportation manager and former professional-football player who tackles traffic delays here.

Spread of Red   It's happening everywhere—red lights are getting longer. And longer. Cities and towns are quietly, albeit reluctantly, giving the green light to longer traffic signals. Through the 1970s, the typical wait at a red light was 45 seconds or so. The wait started growing in the 1980s and 1990s. Today, 90-second spans aren't uncommon, and some are inching near the three-minute mark. Many of these long lights are their longest only at rush hour, but that's little solace to impatient travelers.

Law-abiding drivers stuck at the two-minute-36-second red light at the corner of DuPont Parkway and Frenchtown Road south of Wilmington, Del., take coffee breaks, work on crossword puzzles and, now and then, do some automotive hugging.

"You see couples getting a little more amorous than you'd expect at a traffic light," says Paul Rada, manager of the nearby Quality Inn. "You've got to do something." Mr. Rada himself has read the comics and eaten his lunch sitting at the light, he says. The truly desperate will "jump out of cars and hit the pedestrian button," hoping it will trigger a light change, adds Larry Koczak, a local mail carrier.

Reasons for the lengthening lights vary: More traffic. Wider streets. More lanes dedicated to left turns. And disagreement among traffic engineers who feel just as stuck as the drivers who endure these lights. "It's crept up gradually," says Scott Wainwright, a Montgomery County, Md., official who chairs a national engineering committee on traffic signals.

Longer red lights are intended to help keep busy intersections clear of anything approaching gridlock. In part, that's because several seconds go to waste every time traffic has to start again when red turns to green.

Yielding to Temptation

But longer lights can encourage antsy drivers to race through yellow lights and run reds, causing accidents and pileups that, in addition to risking injury and death, clog traffic even more. Each year, more than 1.8 million crashes occur at intersections, killing more than 7,800 people, the Federal Highway Administration says.

The federal government has launched a special "Stop Red Light Running" initiative, and localities are cracking down with hidden cameras that snap automatic pictures of scofflaws. For

safety's sake, many cities have built more time into stoplight cycles so, for a few seconds, lights at all sides of a given intersection are red—an "all-red" delay designed to clear stragglers and minimize collisions.

All-reds are a source of controversy among traffic engineers. Peter Parsonson, professor of civil engineering at Georgia Tech in Atlanta, says they add an extra margin of safety. But at Marquette University in Milwaukee, David Kuemmel, who is in the department of civil and environmental engineering, argues that the all-red sequence is being used "indiscriminately."

The way to stop frustrated drivers from going through lights is not to make them wait longer, Mr. Kuemmel says. "That's a pet peeve of mine. It's an incentive for motorists to run the red light." The disagreements are further complicated by the lack of conclusive research on the link between light length and crashes.

Many roads are sprouting more left-turn-only lanes, with their own green-arrow turn indicators to keep vehicles that want to turn from backing up forever. But that stretches the light-cycle length. And with extra lanes making streets wider, pedestrians need more time to cross, requiring longer greens to accommodate them.

Even in Las Vegas

"In the old days, when you had two-lane roads, things were different," says Gerry de Camp, a transportation consultant and former manager of the Las Vegas area computer traffic system. Now, he says, "you can have a three-lane left turn, plus a bus lane, and a 12-foot shoulder, yadda, yadda, yadda." Las Vegas has a red light that stretches to two minutes and 45 seconds, although Mr. DeCamp says he tried to keep it from getting any longer. "I wasn't proud of it," he says. "I preferred not to advertise it."

In Northern Virginia, Mr. White, who played cornerback for nine years for the Washington Redskins and other teams, has gone from stopping opposing teams to trying to get vehicles moving. He looks more like a referee on a recent weekday evening as he stands at the Fairfax County intersection, stopwatch in hand, timing a seemingly endless procession of glowing headlights and taillights.

"This is brutal," he says, peering down a line of vehicles that stretches down Fair Lakes Parkway to the horizon, waiting for the light to change. One white sedan waits four minutes and 41 seconds to make a left turn; Mr. White explains why: Southbound traffic on the seven-lane Fairfax County Parkway hasn't eased enough to trigger a change in the computer-controlled light. The light is set to switch automatically only when traffic thins on that road, which typically carries a heavier load than the intersecting one.

"You see any end in sight?" he asks. There is none, for now.

The crossroads appears deceptively rural, with trees and grass at each corner and the nearby pond. But it rests amid

(concluded)

gleaming new condominiums, a shopping mall and an office park, and the roads are lined with signs urging drivers to stop and see the latest housing developments under construction.

Standing on one corner, Mr. White offers running sports-style commentary on the traffic flow, groaning occasionally at a bad move. "The truck is going to run the red!" he shouts. "The truck ran the red!"

A Long Red

He varies the light cycles at different times of the day, and uses the longest cycles only during peak rush-hour traffic. But at those times, he says, he has little choice but to keep the Fair Lakes Parkway red light so as to prevent miles-long backups on the Fairfax County Parkway.

Suddenly, Fair Lakes seems to be clearing up. Cars flow smoothly through the intersection and, by the time the light changes to red, none are left waiting. "This is beautiful!" Mr. White cheers. "We cleared the whole side street out. If it's still that way at 6:30 or so, we can open the champagne."

But Mr. White thinks the light still is too long, and he hopes to do something about it. As part of a computer-plotted plan to smooth traffic flows through 76 local intersections, the longest possible red light here will drop from four minutes and 47 seconds to three minutes and 15 seconds.

It'll still be one of the longest reds anywhere, and Mr. White can't be sure the change will stick. "The goal is to move the traffic," he says. "If it doesn't hold up, we'll change it" back.

Source: Reprinted by permission of Dow Jones & Company, Inc., from "Stopped at a Light? Why Not Read This, You May Have Time" by Anna Wilde Mathews, *The Wall Street Journal,* June 13, 2000, pp. A1 and A12. Copyright © 2000 Dow Jones & Co. Used with permission.

Buffa, Elwood. *Operations Management.* 3rd ed. New York: John Wiley & Sons, 1972.

Griffin, W. *Queuing: Basic Theory and Applications.* Columbus, OH: Grid Publishing, 1978.

Hillier, Frederick S., and Gerald J. Lieberman. *Introduction to Operations Research,* 3rd ed. San Francisco: Holden-Day, 1980.

Katz, K. L., B. M. Larson, and R. C. Larson. "Prescriptions for the Waiting-in-Line Blues: Entertain, Enlighten, and Engage." *Sloan Management Review* 32, no. 2 (Winter 1991), pp. 44–53.

Stevenson, William J. *Introduction to Management Science.* 3rd ed. Burr Ridge, IL: Richard D. Irwin, 1998.

SELECTED BIBLIOGRAPHY AND FURTHER READING

Supplement to
CHAPTER 18

Simulation

LEARNING OBJECTIVES

After completing this supplement, you should be able to:

1 Explain what is meant by the term *simulation.*

2 List some of the reasons for simulation's popularity as a tool for decision making.

3 Explain how and why random numbers are used in simulation.

4 Outline the advantages and limitations of simulation.

5 Describe the alternatives that a manager would reject before choosing simulation as a decision-making tool.

6 Solve typical problems that require the use of simulation.

simulation A descriptive technique that enables a decision maker to evaluate the behavior of a model under various conditions.

Simulation is a descriptive technique in which a model of a process is developed and then experiments are conducted on the model to evaluate its behavior under various conditions. Unlike many of the other models described in the text, simulation is not an optimizing technique. It does not produce a solution per se. Instead, simulation enables decision makers to test *their* solutions on a model that reasonably duplicates a real process; simulation models enable decision makers to experiment with decision alternatives using a *what if* approach.

The use of simulation as a decision-making tool is fairly widespread, and you are undoubtedly familiar with some of the ways it is used. For instance, space engineers simulate space flight in laboratories to permit future astronauts to become accustomed to working in a weightless environment. Similarly, airline pilots often undergo extensive training with simulated landings and takeoffs before being allowed to try the real thing. Many video games are simulations, and universities use management games as a means of simulating business environments. Tire designers evaluate alternative tread designs using machines that simulate conditions that produce tire wear and handling problems.

INTRODUCTION

Simulation has applications across a broad spectrum of operations management problems. In some instances, the simulations are quite modest, while others are rather complex. Their usefulness in all cases depends on the degree to which decision makers are able to successfully answer their *what if* questions.

A list of operations management topics would reveal that most have simulation applications. For instance, simulation is often helpful in product design and testing, facilities layout, line balancing, job design, aggregate planning, testing of alternative inventory policies, scheduling, waiting lines, and project management. Actually, the list is quite long. The books and articles in the bibliography indicate the richness and breadth of those applications and offer the interested reader an opportunity to explore this fascinating and useful subject in more detail.

Generally, analysts use the simulation approach either because optimization techniques are unavailable or because the assumptions required by an optimizing technique are not reasonably satisfied in a given situation. Waiting-line problems are a good example of the latter reason. Although waiting-line problems are pervasive, the rather restrictive assumptions of arrival and service distributions in many cases are simply not met. Very often, analysts will then turn to simulation as a reasonable alternative for obtaining descriptive information about the system in question.

Other reasons for the popularity of simulation include

1. Many situations are too complex to permit development of a mathematical solution; the degree of simplification needed would seriously affect the results. In contrast, simulation models are often able to capture the richness of a situation without sacrificing simplicity, thereby enhancing the decision process.

2. Simulation models are fairly simple to use and understand.

3. Simulation enables the decision maker to conduct experiments on a model that will help in understanding process behavior while avoiding the risks of conducting tests on the model's real-life counterpart.

4. Extensive computer software packages make it easy to use fairly sophisticated models.

5. Simulation can be used for a wide range of situations.

6. There have been numerous successful applications of these techniques.

STEPS IN THE SIMULATION PROCESS

Regardless of the type of simulation involved, certain basic steps are used for all simulation models:

1. Identify the problem and set objectives.

2. Develop the simulation model.

3. Test the model to be sure that it reflects the system being studied.

4. Develop one or more experiments (conditions under which the model's behavior will be examined).

5. Run the simulation and evaluate the results.

6. Repeat steps 4 and 5 until you are satisfied with the results.

The first step in problem solving of any sort is to clearly declare the problem and set objectives that the solution is intended to achieve; simulation is no exception. A clear statement of the objectives can provide not only guidance for model development but also the basis for evaluation of the success or failure of a simulation. In general, the goal of a simulation study is to determine how a system will behave under certain conditions. The more specific a manager is about what he or she is looking for, the better the chances that the simulation model will be designed to accomplish that. Toward that end, the manager must decide on the *scope*

and *level of detail* of the simulation. This indicates the necessary degree of complexity of the model and the information requirements of the study.

The next step is model development. Typically, this involves deciding on the structure of the model and using a computer to carry out the simulations. (For instructional purposes, the examples and problems in this chapter are primarily manual, but in most real-life applications computers are used. This stems from the need for large numbers of runs, the complexity of simulations, and the need for record keeping of results.) Data gathering is a significant part of model development. The amount and type of data needed are a direct function of the scope and level of detail of the simulation. The data are needed for both model development and evaluation. Naturally, the model must be designed to enable evaluation of key decision alternatives.

The validation phase is closely related to model development. Its main purpose is to determine if the model adequately depicts real system performance. An analyst usually accomplishes this by comparing the results of simulation runs with known performance of the system under the same circumstances. If such a comparison cannot be made because, for example, real-life data are difficult or impossible to obtain, an alternative is to employ a test of reasonableness, in which the judgments and opinions of individuals familiar with the system or similar systems are relied on for confirmation that the results are plausible and acceptable. Still another aspect of validation is careful consideration of the assumptions of the model and the values of parameters used in testing the model. Again, the judgments and opinions of those familiar with the real-life system and those who must use the results are essential. Finally, note that model development and model validation go hand in hand: Model deficiencies uncovered during validation prompt model revisions, which lead to the need for further validation efforts and perhaps further revisions.

The fourth step in simulation is designing experiments. Experiments are the essence of a simulation; they help answer the *what if* questions posed in simulation studies. By going through the process, the manager or analyst learns about system behavior.

The fifth step is to run the simulation model. If a simulation model is deterministic and all parameters are known and constant, only a single run will be needed for each *what if* question. But if the model is probabilistic, with parameters subject to random variability, multiple runs will be needed to obtain a clear picture of the results. In this text, probabilistic simulations are the focal point of the discussion, and the comments are limited to them. Probabilistic simulation is essentially a form of random sampling, with each run representing one observation. Consequently, statistical theory can be used to determine appropriate sample sizes. In effect, the larger the degree of variability inherent in simulation results, the greater the number of simulation runs needed to achieve a reasonable level of confidence in the results as true indicators of model behavior.

The last step in the simulation process is to analyze and interpret the results. Interpretation of the results depends to a large extent on the degree to which the simulation model approximates reality; the closer the approximation, the less need to "adjust" the results. Moreover, the closer the approximation of the model to reality, the less the risk inherent in applying the results.

MONTE CARLO SIMULATION

Monte Carlo method Probabilistic simulation technique used when a process has a random component.

random Chance.

IOM

There are many different kinds of simulation techniques. The discussion will focus on probabilistic simulation using the **Monte Carlo method**. The technique gets its name from the famous Mediterranean resort associated with games of chance. The chance element is an important aspect of Monte Carlo simulation, and this approach can be used only when a process has a **random,** or chance, component.

In the Monte Carlo method, an analyst identifies a probability distribution that reflects the random component of the system under study. Random samples taken from this probability distribution are analogous to observations made on the system itself. As the number of observations increases, the results of the simulation will more closely approximate the behavior of

TABLE 18S.1
Random numbers

	1	2	3	4	5	6	7	8	9	10	11	12
1	18	20	84	29	91	73	64	33	15	67	54	07
2	25	19	05	64	26	41	20	09	88	40	73	34
3	73	57	80	35	04	52	81	48	57	61	29	35
4	12	48	37	09	17	63	94	08	28	78	51	23
5	54	92	27	61	58	39	25	16	10	46	87	17
6	96	40	65	75	16	49	03	82	38	33	51	20
7	23	55	93	83	02	19	67	89	80	44	99	72
8	31	96	81	65	60	93	75	64	26	90	18	59
9	45	49	70	10	13	79	32	17	98	63	30	05
10	01	78	32	17	24	54	52	44	28	50	27	68
11	41	62	57	31	90	18	24	15	43	85	31	97
12	22	07	38	72	69	66	14	85	36	71	41	58

the real system, provided an appropriate model has been developed. Sampling is accomplished by the use of random numbers.

The basic steps in the process are as follows:

1. Identify a probability distribution for each random component of the system.
2. Work out an assignment so that intervals of random numbers will correspond to the probability distribution.
3. Obtain the random numbers needed for the study.
4. Interpret the results.

The random numbers used in Monte Carlo simulation can come from any source that exhibits the necessary randomness. Typically, they come from one of two sources: Large studies depend on computer-generated random numbers, and small studies commonly make use of numbers from a table of random digits like the one shown in Table 18S.1. The digits are listed in pairs for convenience, but they can be used singly, in pairs, or in whatever grouping a given problem calls for.

Two important features of the sets of random numbers are essential to simulation. One is that the numbers are uniformly distributed. This means that for any size grouping of digits (e.g., two-digit numbers), every possible outcome (e.g., 34, 89, 00) has the same probability of appearing. The second feature is that there are no discernible patterns in sequences of numbers to enable one to predict numbers further in the sequence (thus the name *random* digits). This feature holds for any sequence of numbers; the numbers can be read across rows and up or down columns.

When using the table, it is important to avoid always starting in the same spot; that would result in the same sequence of numbers each time. Various methods exist for choosing a random starting point. One can use the serial number of a dollar bill to select the row, column, and direction of number selection. Or use rolls of a die. For our purposes, the starting point will be specified in each manual example or problem so that everyone obtains the same results.

The process of simulation will become clearer as we work through some simple problems.

The manager of a machine shop is concerned about machine breakdowns, and wants to simulate breakdowns for a 10-day period. Historical data on breakdowns over the last 100 days are given in the following table:

EXAMPLE S-1

Number of Breakdowns	Frequency
0.............	10
1.............	30
2.............	25
3.............	20
4.............	10
5.............	5
	100

Simulate breakdowns for a 10-day period. Read two-digit random numbers from Table 18S.1, starting at the top of column 1 and reading down.

SOLUTION

a. Develop cumulative frequencies for breakdowns:

(1) Convert frequencies into relative frequencies by dividing each frequency by the sum of the frequencies. Thus, 10 becomes $10/100 = .10$, 30 becomes $30/100 = .30$, and so on.

(2) Develop cumulative frequencies by successive summing. The results are shown in the following table:

Number of Breakdowns	Frequency	Relative Frequency	Cumulative Frequency
0	10	.10	.10
1	30	.30	.40
2	25	.25	.65
3	20	.20	.85
4	10	.10	.95
5	5	.05	1.00
	100	1.00	

b. Assign random-number intervals to correspond with the cumulative frequencies for breakdowns. (*Note:* Use two-digit numbers because the frequencies are given to two decimal places.) You want a 10 percent probability of obtaining the event "0 breakdowns" in our simulation. Therefore, you must designate 10 percent of the possible random numbers as corresponding to that event. There are 100 two-digit numbers, so you can assign the 10 numbers 01 to 10 to that event.

Similarly, assign the numbers 11 to 40 to "one breakdown," 41 to 65 to "two breakdowns," 66 to 85 to "three breakdowns," 86 to 95 to "4 breakdowns" and 96 to 00 to five breakdowns.

Number of Breakdowns	Frequency	Probability	Cumulative Probability	Corresponding Random Numbers
0	10	.10	10	01 to 10
1	30	.30	.40	11 to 40
2	25	.25	.65	41 to 65
3	20	.20	.85	66 to 85
4	10	.10	.95	86 to 95
5	5	.05	1.00	96 to 00
	100	1.00		

c. Obtain the random numbers from Table 18S.1, column 1, as specified in the problem:

18   25   73   12   54   96   23   31   45   01

d. Convert the random numbers into numbers of breakdowns:

18 falls in the interval 11 to 40 and corresponds, therefore, to one breakdown on day 1.

25 falls in the interval 11 to 40, which corresponds to one breakdown on day 2.

73 corresponds to three breakdowns on day 3.

12 corresponds to one breakdown on day 4.

54 corresponds to two breakdowns on day 5.

96 corresponds to five breakdowns on day 6.

23 corresponds to one breakdown on day 7.

31 corresponds to one breakdown on day 8.

45 corresponds to two breakdowns on day 9.

01 corresponds to no breakdowns on day 10.

The following table summarizes these results:

Day	Random Number	Simulated Number of Breakdowns
1	18	1
2	25	1
3	73	3
4	12	1
5	54	2
6	96	5
7	23	1
8	31	1
9	45	2
10	01	0
		17

The mean number of breakdowns for this 10-period simulation is 17/10 = 1.7 breakdowns per day. Compare this to the *expected* number of breakdowns based on the historical data:

$$0(.10) + 1(.30) + 2(.25) + 3(.20) + 4(.10) + 5(.05) = 2.05 \text{ per day}$$

Several points are worth noting:

1. This simple example is intended to illustrate the basic concept of Monte Carlo simulation. If our only goal were to estimate the average number of breakdowns, we would not have to simulate; we could base the estimate on the historical data alone.

2. The simulation should be viewed as a *sample*; it is quite likely that additional runs of 10 numbers would produce different means.

3. Because of the variability inherent in the results of small samples, it would be unwise to attempt to draw any firm conclusions from them; in an actual study, much larger sample sizes would be used.

In some cases, it is helpful to construct a flowchart that describes a simulation, especially if the simulation will involve periodic updating of system values (e.g., amount of inventory on hand), as illustrated in Example S–2.

The Excel spreadsheet formulation for this problem is shown below. Note that the alignment of values in columns B, C, and F must be exactly as shown.

	A	B	C	D	E	F	G
1	EXAMPLE S-1						
2							
3	Probability	Cumulative probability	Number of breakdowns		Day	Random number	Simulated demand
4	0	0	0		1	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
5	0.1	0.1	1		2	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
6	0.3	0.4	2		3	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
7	0.25	0.65	3		4	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
8	0.2	0.85	4		5	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
9	0.1	0.95	5		6	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
10	0.05	1			7	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
11					8	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
12					9	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
13					10	=RAND()	=VLOOKUP(F4,$B$4:$C$10,2,1)
14						SUM	=SUM(G4:G13)
15						AVG	=AVERAGE(G4:G13)

The simulation results are shown in the following screen. Use key F4 to do a simulation or another simulation.

	A	B	C	D	E	F	G
1	EXAMPLE S-1						
2							
3	Probability	Cumulative probability	Number of breakdowns		Day	Random number	Simulated demand
4	0	0	0		1	0.22433	1
5	0.1	0.1	1		2	0.39772	1
6	0.3	0.4	2		3	0.20817	1
7	0.25	0.65	3		4	0.631	2
8	0.2	0.85	4		5	0.28439	1
9	0.1	0.95	5		6	0.16299	1
10	0.05	1			7	0.9581	5
11					8	0.66967	3
12					9	0.07676	0
13					10	0.06033	0
14						SUM	15
15						AVG	1.5

EXAMPLE S-2

The manager of a small truck dealership wants to acquire some insight into how a proposed policy for reordering trucks might affect order frequency. Under the new policy, two trucks are to be ordered whenever the number of trucks on hand is five or fewer. Due to the nearness of the dealer to the home office, orders can be filled overnight. According to the dealer's records, the probability distribution for daily demand is

Demand, x	P(x)
0........	.50
1........	.40
2........	.10

a. Construct a flowchart that describes a 10-day simulation.

b. Use two-digit random numbers from Table 18S.1, column 11, reading down. Assume a beginning inventory of seven trucks.

a.

Start
$I = 7$
$D = 1$

x = Demand
D = Day
I = Amount of inventory

$I \leq 5?$ Yes

No

$D = 10?$ No

Yes
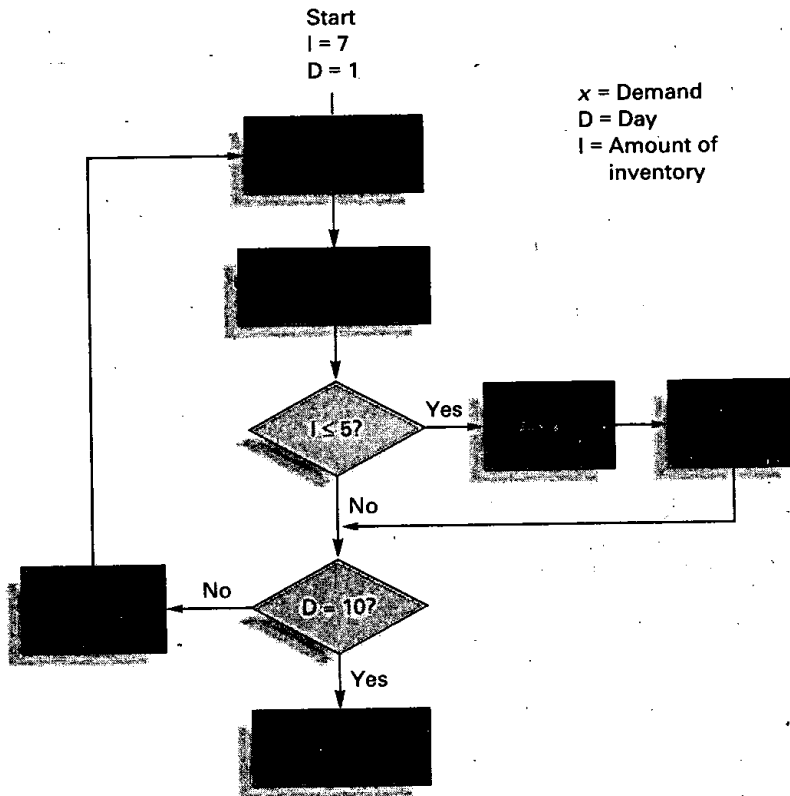
b. (1) Specify random number ranges for demand:

x	P(x)	Cumulative P(x)	Ranges
0	.50	.50	01–50
1	.40	.90	51–90
2	.10	1.00	91–00

(2) Obtain random numbers, convert to demand, update inventory accordingly, and re-order when necessary:

Day	Random Number	Demand, x	Beginning Inventory	Ending Inventory
1	54	1	7	6
2	73	1	6	5 (reorder 2; new beginning inventory = 5 + 2)
3	29	0	7	7
4	51	1	7	6
5	87	1	6	5 (reorder 2; new beginning inventory = 5 + 2)
6	51	1	7	6

7 99	2	6	4 (reorder 2; new beginning inventory = 4 + 2)
8 18	0	6	6
9 30	0	6	6
10........ 27	0	6	6

Simulating Theoretical Distributions

In many instances, a simulation will involve the use of theoretical distributions. Among the most frequently encountered theoretical distributions are the Poisson, normal, and exponential distributions. Being able to simulate these distributions will greatly enhance your knowledge and appreciation of simulation.

Simulation of a Poisson distribution requires the mean of the distribution. Knowledge of the mean enables one to obtain cumulative probabilities for the distribution from Appendix B, Table C; these, in turn, provide the basis for random-number assignments. You can use Table 18S.1 to obtain random numbers; you must read three-digit random numbers from Table 18S.1 to achieve correspondence. Example S–3 illustrates these concepts.

EXAMPLE S-3

The number of lost-time accidents at a large plant has been determined from historical records to be two per day. Moreover, it has been determined that this accident rate can be well approximated by a Poisson distribution that has a mean of 2.0. Simulate five days of accident experience for the plant. Read random numbers from columns 1 and 2 of Table 18S.1.

SOLUTION

First obtain the cumulative distribution from Appendix B, Table C for a mean of 2.0, and make the range assignments:

x	Cumulative Probability	Random Number Ranges
0	.135	001 to 135
1	.406	136 to 406
2	.677	407 to 677
3	.857	678 to 857
4	.947	858 to 947
5	.983	948 to 983
6	.995	984 to 995
7	.999	996 to 999
8	1.000	000

Next obtain three-digit numbers from Table 18S.1. Reading from columns 1 and 2 as instructed, you find 182, 251, 735, 124, and 549.

Finally, convert the random numbers into number of lost-time accidents using the established set of ranges. Since 182 falls in the second range, it corresponds to one accident on day 1. The second random number, 251, falls in the same range, indicating one accident on day 2. The number 735 falls between 678 and 857, which corresponds to three accidents on day 3; 124 corresponds to 0 accidents on day 4; and 549 corresponds to two accidents on day 5.

The normal distribution can be important in many problems. There are a number of ways to simulate a normal distribution, but perhaps the simplest is to use a table of normally distributed random numbers, such as Table 18S.2. The basis of the table is a normal distribution with a mean of 0 and a standard deviation of 1.00. Like all such tables, the numbers are arranged randomly, so that when they are read in any sequence they exhibit "randomness." In order to use the table, we must have the parameters of a normal distribution (i.e., its mean and standard deviation) in mind. Numbers obtained from the random number table can then be converted to "actual" values by multiplying the standard deviation by the random number and adding this amount to the mean. That is:

TABLE 18S.2
Normally distributed random numbers

	1	2	3	4	5	6	7	8	9	10
1	1.46	-0.09	-0.59	0.19	-0.52	-1.82	0.53	-1.12	1.36	-0.44
2	-1.05	0.56	-0.67	-0.16	1.39	-1.21	0.45	-0.62	-0.95	0.27
3	0.15	-0.02	0.41	-0.09	-0.61	-0.18	-0.63	-1.20	0.27	-0.50
4	0.81	1.87	0.51	0.33	-0.32	1.19	2.18	-2.17	1.10	0.70
5	0.74	-0.44	1.53	-1.76	0.01	0.47	0.07	0.22	-0.59	-1.03
6	-0.39	0.35	-0.37	-0.52	-1.14	0.27	-1.78	0.43	1.15	-0.31
7	0.45	0.23	0.26	-0.31	-0.19	-0.03	-0.92	0.38	-0.04	0.16
8	2.40	0.38	-0.15	-1.04	-0.76	1.12	-0.37	-0.71	-1.11	0.25
9	0.59	-0.70	-0.04	0.12	1.60	0.34	0.05	-0.26	0.41	0.80
10	-0.06	0.83	-1.60	-0.28	0.28	-0.15	0.73	-0.13	-0.75	-1.49

Simulated value = Mean + Random number × Standard deviation $\qquad$ (18S–1)

In effect, the random number equates to a normal z value, which indicates how far a particular value is above or below the distribution mean.

It has been determined that the time required to perform a certain task can be described by a normal distribution that has a mean of 30 minutes and a standard deviation of 4 minutes. Simulate times for three jobs using the first three values in column 1 of Table 18S.2.

EXAMPLE S-4

SOLUTION

The first three values are 1.46, −1.05, and 0.15. The simulated values are

For	1.46:	30 + 1.46(4) = 35.84 minutes.
For	−1.05:	30 − 1.05(4) = 25.80 minutes.
For	0.15:	30 + 0.15(4) = 30.60 minutes.

It is important to recognize that Example S–4 involves a continuous variable, whereas the previous examples involved discrete variables. (Remember that discrete variables typically take on only integer values, whereas continuous variables can take on integer and noninteger values.) Whenever possible, a model of a continuous variable should be able to simulate noninteger values as well as integer values.

Another continuous type of distribution we can consider is the uniform distribution, in which values may occur anywhere over a continuous range between two extremes, a and b, as illustrated in Figure 18S.1.

Simulated value = $a + (b - a)$(Random number as a percentage) $\qquad$ (18S–2)

Converting the random number to a percentage simply involves placing a decimal point to the left of the number. For example, 77 becomes .77.

Example S–5 illustrates a simulation involving a uniform distribution.

A third continuous distribution is the exponential distribution. We will concern ourselves with simulating values of negative exponential distributions, as portrayed in Figure 18S.2.
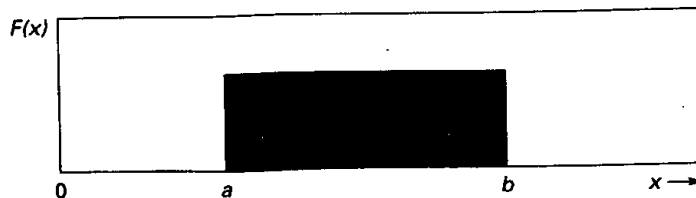

FIGURE 18S.1

A uniform distribution

EXAMPLE S-5

Job times vary uniformly between 10 and 15 minutes. Use Table 18S.1 to simulate job times for four jobs. Read numbers from column 9, going down.
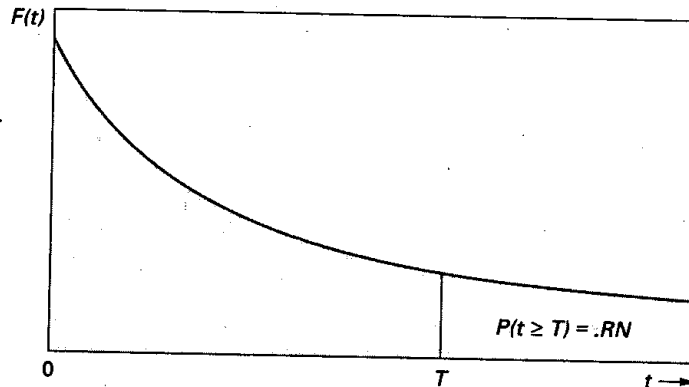
SOLUTION ————

$$a = 10 \text{ minutes}, \quad b = 15 \text{ minutes}, \quad b - a = 5 \text{ minutes}$$

a. Obtain the random numbers: 15, 88, 57, and 28.

b. Convert to simulated values:

Random Number	Computation		Simulated Value (minutes)
15	10 + 5(.15)	=	10.75
88	10 + 5(.88)	=	14.40
57	10 + 5(.57)	=	12.85
28	10 + 5(.28)	=	11.40

FIGURE 18S.2

A negative exponential distribution

With a negative exponential distribution, the probability is fairly high that the random variable will assume a value close to zero. Moreover, the probability decreases as the specified value of the random variable increases. The probability that a random variable will take on a value greater than some specified value T, given that the variable can be described by an exponential distribution with a mean equal to $1/\lambda$, is given by the equation

$$P(t \geq T) = e^{-\lambda t} \tag{18S-3}$$

To simulate exponential values, we obtain a random number, place a decimal point to the left of it, set this equal to the probability $P(T)$, and solve Formula 18S–3 for t. The result is a simulated value from an exponential distribution with a mean of $1/\lambda$.

We can obtain an expression for t by taking the natural logarithm of both sides of the equation. Thus, with $P(T) = .RN$ (for *random number*), we have

$$\ln(.RN) = \ln(e^{-\lambda t})$$

The natural logarithm of a power of e is equal to the power itself, so

$$\ln(.RN) = \ln(e^{-\lambda t}) = -\lambda t$$

Then

$$t = -\frac{1}{\lambda} \ln(.RN) \tag{18S-4}$$

This concept is illustrated in Figure 18S.2.

Values of random numbers can be obtained using Table 18S.1, as demonstrated in the following example.

Times between breakdowns of a certain type of equipment can be described by an exponential distribution with a mean of five hours. Simulate the time between two breakdowns. Read two-digit random numbers from column 3 of Table 18S.1.

EXAMPLE S-6

SOLUTION

The mean, $1/\lambda$, is 5 hours. The random numbers are 84 and 05. Using Formula 18S-4, the simulated times are

For 84: $t = -5[\ln(.84)] = -5[-0.1744] = 0.872$ hour

For 05: $t = -5[\ln(.05)] = -5[-2.9957] = 14.979$ hours

Note that the smaller the value of the random number, the larger the simulated value of t.

COMPUTER SIMULATION

Although the emphasis in this supplement has been on manual simulation in order to convey the main concepts, most real-life simulations involve the use of a computer. Computers offer relatively quick and easy means of obtaining results. Many real-life applications involve fairly complex simulation models that have substantial record-keeping requirements for which computers offer a practical solution.

Over the years, programmers have developed a number of simulation languages that make the task of writing simulation programs much simpler. Some of the general-purpose languages are SIMSCRIPT II.5, GPSS/H, GPSS/PC, and RESQ. In addition, a number of other simulation packages are available, some of which have very narrow focuses that relate to queuing or network problems. Most of the simulation packages have certain features in common. For example, they generally provide for random number generation from a variety of statistical distributions, as well as collection and tabulation of simulation results and time keeping.

It should be noted that some managers prefer to write their own simulations or have a member of their staff do so, using spreadsheet software such as MS Excel rather than a simulation language, as illustrated on page 820. In cases where simulation is used infrequently, it can be more practical to use that approach rather than go through the time and effort required to use a specialty language. For simple problems, the slight inefficiency related to the use of a standard language is not usually an important consideration.

ADVANTAGES AND LIMITATIONS OF USING SIMULATIONS

Among the main advantages of simulation are these:

1. It lends itself to problems that are difficult or impossible to solve mathematically.
2. It permits an analyst to experiment with system behavior while avoiding possible risks inherent in experimenting with the actual system.
3. It compresses time so that managers can quickly discern long-term effects.
4. It can serve as a valuable tool for training decision makers by building up their experience and understanding of system behavior under a wide range of conditions.

Certain limitations are also associated with simulation. Chief among these are

1. Simulation does not produce an optimum solution; it merely indicates an *approximate* behavior for a given set of inputs. There are two reasons for this:
   a. By design, there is inherent randomness (i.e., random numbers) in simulation.
   b. Simulations are based on models, and models are only approximations of reality.
2. For large-scale simulation, it can require considerable effort to develop a suitable model as well as considerable computer time to obtain simulations.

FIGURE 18S.3

Define the problem

Is an intuitive solution acceptable? — Yes → Use the intuitive solution

No

Is an analytical model available? — Yes → Use the model

No

Can an analytical model be developed? — Yes → Use the model

No

Is simulation economically justified? — Yes → Use simulation

No

Rely on judgment and experience
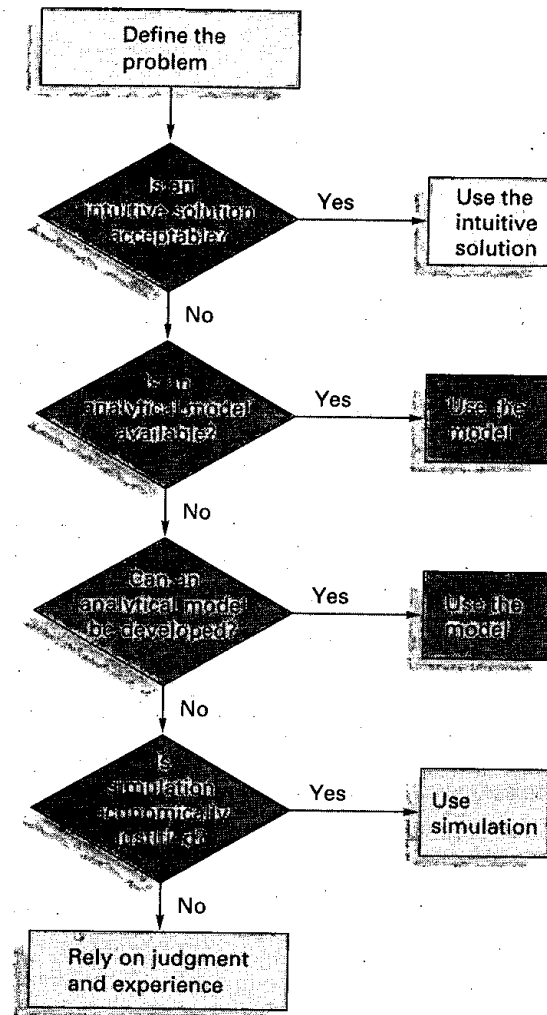
Because simulation produces an approximate answer rather than an exact solution, and because of the cost of running a simulation study, simulation is not usually the first choice of a decision maker. Instead, depending on the complexity of the situation, intuitive or analytical methods should first be investigated. In simple cases, an intuitive solution very often is acceptable. In more complex cases, an analytical solution is preferable, assuming an appropriate technique is available. If not, it may be possible to develop an analytical model that could be used to generate a solution. If these measures do not suffice, simulation becomes the next logical possibility. Of course, if that is not economically justifiable, the decision maker will have to rely on judgment and experience; in effect, after reevaluating all of the alternatives, the decision maker may revert to an intuitive solution, even though initially that approach did not appear acceptable. Figure 18S.3 outlines this process.

KEY TERMS
Monte Carlo method, 816
random, 816

simulation, 814

Problem 1

The number of customers who arrive at a transmission repair shop can be described by a Poisson distribution that has a mean of three per hour. Assuming the distribution holds for an entire eight-hour day, simulate customer arrivals for the first four hours of a day. Read random numbers from Table 18S.1, columns 4 and 5, going down.

Solution

a. Obtain cumulative probabilities of the Poisson distribution from Appendix B, Table C for the mean specified. (Those values are given below for convenience.) Determine the random number ranges.

x	Cumulative Probability
0.......	.050
1.......	.199
2.......	.423
3.......	.647
4.......	.815
5.......	.916
6.......	.966
7.......	.988
8.......	.996
9.......	.999
10.......	1.000

b. Obtain the random numbers: 299, 642, 350, and 091. (*Note:* Three-digit numbers are needed because the probabilities are given to three decimal places.)

c. Convert the random numbers to numbers of arrivals. Note where each number falls in the random number range list. For instance, 299 falls between 199 and 423. Interpret this to mean two customers arrive in the first hour. Similarly, 642 is interpreted to mean three customers arrive in the second hour, 350 implies two customers in the third hour, and 091 implies one customer in the fourth hour.

In sum, the number of customers per hour for the four-hour simulation is

Hour	Number of Arrivals
1...........	2
2...........	3
3...........	2
4...........	1

Problem 2

Jobs arrive at a workstation at fixed intervals of one hour. Processing time is approximately normal and has a mean of 56 minutes per job and a standard deviation of 4 minutes per job. Using the fifth row of the table of normally distributed random numbers (Table 18S.2), simulate the processing times for four jobs, and determine the amount of operator idle time and job waiting time. Assume the first job arrives at time = 0.

Solution

a. Obtain the random numbers from the table: 0.74, −0.44, 1.53, and −1.76.

b. Convert the random numbers to simulated processing times:

Random Number	Computation	Simulated Time
0.74	56 + 4(0.74) =	58.96
−0.44	56 + 4(−0.44) =	54.24
1.53	56 + 4(1.53) =	62.12
−1.76	56 + 4(−1.76) =	48.96

Note that three of the times are less than the interarrival times for the jobs, meaning the operator may be idle after those three jobs. One time exceeds the one-hour interval, so the next job must wait, and possibly the job following it if the waiting plus processing time exceeds 60 minutes.

c. Compute waiting and idle times:

Job Number	Arrives at	Processing Time, t (minutes)	$60 - t$ Operator Idle (minutes)	$t - 60$ Next Job Waits (minutes)
1	0	58.96	1.04	—
2	60	54.24	5.76	—
3	120	62.12	—	2.12
4	180	48.96	8.92*	—
			15.72	2.12

*$60 - 2.12 - 48.96 = 8.92.$

Problem 3

The time between mechanics' requests for tools in a large plant is normally distributed with a mean of 10 minutes and a standard deviation of 1 minute. The time to fill requests is also normal with a mean of 9 minutes per request and a standard deviation of 1 minute. Mechanics' waiting time represents a cost of $2 per minute, and servers represent a cost of $1 per minute. Simulate arrivals for the first nine mechanic requests and their service times, and determine the mechanics' waiting time, assuming one server. Would it be economical to add another server? Explain. Use Table 18S.2, column 8 for requests and column 9 for service.

Solution

a. Obtain random numbers and convert to times [see columns (a) and (b) in the following table for requests and columns (f) and (g) for service].

CUSTOMER ARRIVALS					SERVICE		
(a) Random Number	(b) Time between Arrivals	(c) Cumulative Arrival Time	(d) ($e - c$) Customer Waiting Time	(e) Service Begins	(f) Random Number	(g) Service Time	(h) ($e + g$) Service Ends
−1.12	8.88	8.88	0.00	8.88	1.36	10.36	19.24
−0.62	9.38	18.26	0.98	19.24	−0.95	8.05	27.29
−1.20	8.80	27.06	0.23	27.29	0.27	9.27	36.56
−2.17	7.83	34.89	1.67	36.56	1.10	10.10	46.66
0.22	10.22	45.11	1.55	46.66	−0.59	8.41	55.07
0.43	10.43	55.54	0.00	55.54	1.15	10.15	65.69
0.38	10.38	65.92	0.00	65.92	−0.04	8.96	74.88
−0.71	9.29	75.21	0.00	75.21	−1.11	7.89	83.10
−0.26	9.74	84.95	0.00	84.95	0.41	9.41	94.36
			4.43				

b. Determine arrival times [column (c)] by successive adding to times between arrivals in column (b).

c. Use arrival times for service start *unless service is still in progress on a previous request.* In that case, determine how long the arrival must wait ($e - c$). Column (e) values are the sum of starting time and service time [column (g)], which is the time service ends [column (h)]. Thus, service on each new request begins [column (e)] at the same time that service on the previous request ends [column (h)].

d. The simulation and resulting waiting times for the first nine arrivals are shown in the table. Total waiting time is 4.43 minutes.

e. The total cost for the 94.36 minutes (end of service on the ninth request) of the simulation is

Waiting cost: 4.43 minutes at $2 per minute = $ 8.86

Server cost: 94.36 minutes at $1 per minutes =   94.36
                                                     ――――――
                                                     $103.22

f. Usually, a second simulation with two servers would be needed (but with the same arrival times so that the results are comparable). However, in this case it is apparent that a second server would increase server cost by about $94 but could not eliminate more than $8.86 of waiting cost. Hence, the second server would not be justified.

================= **DISCUSSION AND REVIEW QUESTIONS**

1. What is a simulation?
2. What are some of the primary reasons for the widespread use of simulation techniques in practice?
3. What are some of the ways managers can use simulation?
4. What role do random numbers play in Monte Carlo simulations?
5. Respond to the following comment: "I ran the simulation several times, and each run gave me a different result. Therefore, the technique does not seem to be useful. I need answers!"
6. List the main advantages of simulation.
7. What are some of the limitations of simulation as a tool for decision making?

================= **PROBLEMS**

1. The number of jobs received by a small shop is to be simulated for an eight-day period. The shop manager has collected the following data:

Number of Jobs	Frequency
2 or less	0
3	10
4	50
5	80
6	40
7	16
8	4
9 or more	0
	200

   Use the third column of Table 18S.1 and read two-digit numbers, going down. Determine the average number of jobs per day for the eight-day simulation period.

2. Jack M. sells insurance on a part-time basis. His records on the number of policies sold per week over a 50-week period are

Number Sold	Frequency
0	8
1	15
2	17
3	7
4	3
	50

   Simulate three five-day periods. Use Table 18S.1, column 6 for the first simulation, column 7 for the second, and column 8 for the third. In each case, read two-digit numbers, beginning at the *bottom* of the column and going *up*. For each simulation, determine the percentage of days on which two or more policies are sold.

3. After a careful study of requests for a special tool at a large tool crib, an analyst concluded that demand for the tool can be adequately described by a Poisson distribution with a mean of two requests

per day. Simulate demand for a 12-working-day period for this tool using Table 18S.1. Read three-digit numbers from columns 5 and 6 combined, starting at the *top* and reading *down* (e.g., 917, 264, 045).

4. The number of lost-time accidents at a logging firm can be described using a Poisson distribution that has a mean of four accidents per month. Using the last two columns of Table 18S.1 (e.g. 540, 733, 293), simulate accidents for a 12-month period.

5. The time a physician spends with patients can be modeled using a normal distribution that has a mean of 20 minutes and a standard deviation of 2 minutes. Using the table of normally distributed random numbers (Table 18S.2), simulate the times the doctor might spend with the next seven patients. Use column 4 of the table; start at the *bottom* of the column and read *up*.

6. Jobs are delivered to a workstation at random intervals. The time between job arrivals tends to be normally distributed with a mean of 15 minutes and a standard deviation of 1 minute. Job processing time is also normally distributed with a mean of 14 minutes per job and a standard deviation of 2 minutes.

    a. Using Table 18S.2, simulate the arrival and processing of five jobs. Use column 4 of the table for job arrival times and column 3 for processing times. Start each column at row 4. Find the total times jobs wait for processing.

    b. The company is considering the use of new equipment that would result in processing time that is normal with a mean of 13 minutes and a standard deviation of 1 minute. Job waiting represents a cost of $3 per minute, and the new equipment would represent an additional cost of $.50 per minute. Would the equipment be cost justified? (*Note:* Use the same arrival times and the same random numbers for processing times.)

7. Daily usage of sugar in a small bakery can be described by a uniform distribution with endpoints of 30 pounds and 50 pounds. Assuming usage is independent of the day of the week, simulate daily usage for a 10-day period. Read four-digit numbers from Table 18S.1, columns 5 and 6, going *up* from the bottom.

8. Weekly usage of spare parts for a specialized machine can be described by a Poisson distribution with a mean of 2.8 parts per week. Lead time to replenish the supply of spare parts is one week (constant). Simulate the total usage of parts during lead time 10 times, and then determine the frequency of lead time demands (i.e., what percentage of times was the demand equal to 2, 3, 4, etc.?). Read four-digit numbers from Table 18S.1, columns 8 and 9, going down.

9. (*Computer exercise.*) Repeat Problem 8 for 100 lead time periods.

10. A repair shop breaks an average of .6 tool per day. The average number of days required to obtain replacements is six. (Parts are delivered by mail once each day.) Both breakages and delivery times can be described by Poisson distributions. Tools are reordered whenever three or more must be replaced.

    a. Draw a flowchart to describe this process.

    b. Simulate breakage and ordering for a 12-day period. Read three-digit numbers from Table 18S.1, columns 5 and 6, going down (e.g., 917, 264), for tool breakage, and columns 7 and 8, going down (e.g., 643, 200), for delivery time. Assume zero tools in inventory to start.

11. (*Computer exercise.*) Repeat Problem 10 for 150 days.

12. Customers arrive randomly at a catalog department of a large store. The time between arrivals varies uniformly between 10 and 20 minutes. Service time is normal with a mean of 15 minutes and a standard deviation of 2 minutes.

    a. Simulate processing and waiting times for nine customers. Read three-digit numbers going down columns 9 and 10 of Table 18S.1 for arrivals (e.g., 156, 884, 576). Use column 8, Table 18S.2, for processing time.

    b. If management can reduce the range of arrival times to between 13 and 17 minutes, what would the impact be on customer waiting times? (Use the same service times and the same random numbers for arrival times from part *a.*) Round arrival times to two decimal places.

13. Probabilities have been determined for the movements of the ball in a pinball game. These are shown in the accompanying table, along with the points awarded if the ball strikes a given position. Simulate the paths of three balls, and compute the number of points awarded for each ball. Use column 1 of Table 18S.1 for the first ball, column 2 for the second ball, and column 3 for the third ball, reading down the columns.
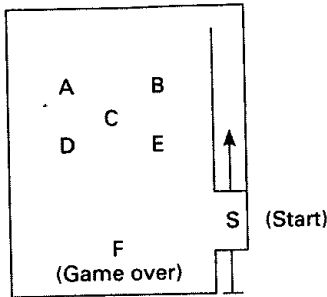
Position	Points
A........	400
B........	300
C........	200
D........	100
E........	50
F........	20

```
 _____
|   A     B      | |
|      C         | |
|   D     E      |↑|
|                | |
|             S  |_| (Start)
|      F         | |
| (Game over)    |_|
|_____|
```

Path	Probability	Path	Probability	Path	Probability
S to A	.30	B to A	.30	D to A	.10
B	.30	B	.20	B	.05
C	.25	C	.20	C	.25
D	.10	D	.15	D	.20
E	.04	E	.10	E	.25
F	.01	F	.05	F	.15
A to A	.25	C to A	.10	E to A	.05
B	.25	B	.15	B	.10
C	.25	C	.20	C	.15
D	.10	D	.15	D	.15
E	.10	E	.25	E	.20
F	.05	F	.15	F	.35

(*Hint:* Construct a cumulative probability distribution for each of the six cases. Each new random number indicates the next position to move to. Each ball must start at S.)

14. Repeat problem 13, using columns 1, 2, and 3 for the next three balls, reading up from the bottom.

15. An analyst found that the length of telephone conversations in an office could be described by an exponential distribution with a mean of four minutes. Reading two-digit random numbers from Table 18S.1, column 6, simulate the length of five calls and compute the simulated average time. Why is the simulated average different from the mean of four minutes?

16. The length of time between calls for service of a certain piece of equipment can be described by an exponential distribution with a mean of 40 minutes. Service time can be described by a normal distribution with a mean of eight minutes and a standard deviation of two minutes. Simulate the time until the first breakdown and the times between three breakdowns, as well as the four service times. For breakdowns, read two-digit numbers from Table 18S.1, column 7; for service times, read numbers from Table 18S.2, column 8.

17. The number of jobs per day that a repair shop gets can be described by a Poisson distribution with a mean of 3.0. Repair time per job can be described by an exponential distribution with a mean of six hours. Simulate the number of jobs received for a four-day period and the repair time for each job. What is the simulated total repair time per day? For number of jobs received, read three-digit numbers from Table 18S.1, row 1; for repair times, read two-digit numbers from row 2.

18. A service operation consists of three steps. The first step can be described by a uniform distribution that ranges between five and nine minutes. The second step can be described by a normal distribution with a mean of seven minutes and a standard deviation of one minute, and the third step can be described by an exponential distribution with a mean of five minutes. Simulate three cycles using two-digit numbers from Table 18S.1, row 4 for step 1; Table 18S.2, row 6 for step 2; and two-digit numbers from column 4 of Table 18S.1 for step 3. Determine the simulated time for each of the three cycles.

19. A project consists of five major activities, as illustrated in the accompanying diagram. Activity times are normally distributed with means and standard deviations as shown in the following table. Note that there are two paths through the project: 1-2-3-5 and 1-2-4-5. Project duration is defined as the largest sum of times along a path. Simulate 12 times for each activity. Use columns 1 and 2 of Table 18S.2 for

activity 1-2, columns 3 and 4 for activity 2-3, columns 5 and 6 for activity 2-4, columns 7 and 8 for activity 3-5, and columns 9 and 10 for activity 4-5. Determine the project duration for each of the 12 sets, and prepare a frequency distribution of project duration. Use categories of 25 to less than 30, 30 to less than 35, 35 to less than 40, 40 to less than 45, and 45 or more. Determine the proportion of time that a simulated duration of less than 40 days occurred. How might this information be used?

Activity	Mean (days)	Standard Deviation (days)
1-2	10	2
2-3	12	2
2-4	15	3
3-5	14	2
4-5	8	1
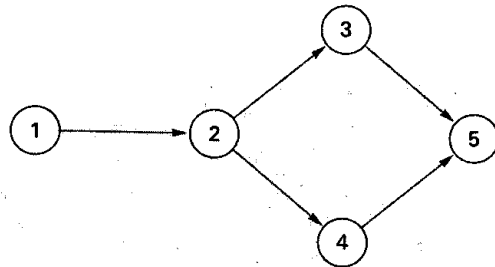
Coquille Refinery Corporation

CASE

The Coquille Refinery Corporation is contemplating building a crude-oil storage and docking facility on the southern coast of France. They plan to import crude oil by ship from the Middle East and distribute this crude oil by pipeline to refineries in the area.

The construction of this facility represents a substantial capital investment. Furthermore, the cost of such a facility is principally determined by its crude-oil storage capacity. You have been asked to study the problem and recommend an appropriate storage capacity; bear in mind that too large a capacity represents an unnecessary expense, but too small a capacity will result in costly later additions to the facility.

A long-term contract has been made with the Middle East supplier to furnish an average daily supply of 300,000 barrels of crude oil. Because its fleet of ships consists of 200,000-barrel tankers, the supplier expects that the arrival of its tankers will follow the distribution below:

Tanker Arrivals per Day	Probability
0	0.1
1	0.5
2	0.2
3	0.2

A review of past production records of refineries in the area suggests the following distribution of crude-oil demand per day:

Barrels per Day	Probability
100,000	0.1
200,000	0.2
300,000	0.3
400,000	0.4

Questions

1. Consider the following issues before you simulate:
   a. What is the expected daily demand for crude oil? Why must this be so?
   b. What assumptions concerning the timing of crude-oil receipts and deliveries would require the greatest oil storage capacity?
   c. What assumption concerning receipts and deliveries would require the least oil storage capacity?
   d. Give a reason based on systems-analysis considerations why back orders should be filled from the next day's receipts rather than considered as lost sales.

(continued)

(concluded)

2. Develop a Monte Carlo simulation model for Coquille that will generate information useful for resolving the storage capacity problem and simulate 10 days of capacity.
3. Assume a computer simulated 10,000 days of activity under the assumptions listed in question 1b. From the results of the simulation, the following distribution of oil in storage, after a day's receipts, is determined. (Note: Negative figures represent back orders.)

Oil in Storage, in Thousands	Probability
−300	0.01
−200	0.04
−100	0.06
0	0.07
⋮	⋮
1,000	0.09
1,100	0.08
1,200	0.07
1,300	0.05
1,400	0.03
1,500	0.02
	1.00

a. What level of crude-oil safety stock should be carried to ensure 95 percent protection against stock-outs (1 stock-out in 20 days)?
b. If Coquille decides to use the above safety stock, what should be the oil storage capacity to ensure 90 percent protection against overruns (i.e., sufficient capacity to completely accommodate receipts 9 days out of 10)?
4. Note two ways one might determine if the simulation run length of 10,000 days is adequate.
5. Make a list of cost factors that would influence the final selection of oil storage capacity.

Source: James A. Fitzsimmons and Robert S. Sullivan, *Service Operations Management* (New York: McGraw-Hill, 1982), pp. 78–80. Reprinted by permission.

Evans, James R., and David L. Olson. *Introduction to Simulation and Risk Analysis.* Upper Saddle River, NJ: Prentice Hall, 1998.

Harrell, C., and K. Tumay. *Simulation Made Easy: A Manager's Guide.* Norcross, GA: Industrial Engineering and Management Press, 1995.

Hillier, Frederick S., and Mark S. Hillier. *Introduction to Management Science.* 2/e. New York: McGraw-Hill/Irwin, 2003.

Law, Averill, and David Kelton. *Simulation Modeling and Analysis.* 3rd ed. New York: McGraw-Hill/Irwin, 2000.

SELECTED BIBLIOGRAPHY AND FURTHER READING