

Designers frequently use computer models to assist them in carrying out their activities. Examples include finite element models for many aspects of structural and mechanical design, electrical circuit simulators for integrated circuit design, factory or enterprise-level models for scheduling and capacity planning or supply chain management, and computer models of complex chemical processes. Statistically designed experiments can be applied to these models just as easily and successfully as they can to actual physical systems and will result in reduced development lead time and better designs.

### EXAMPLE 1.5 Formulating a Product

A biochemist is formulating a diagnostic product to detect the presence of a certain disease. The product is a mixture of biological materials, chemical reagents, and other materials that when combined with human blood react to provide a diagnostic indication. The type of experiment used here is a **mixture experiment**, because various ingredients that are combined to form the diagnostic make up 100 percent of the mixture composition (on a volume, weight, or

mole ratio basis), and the response is a function of the mixture proportions that are present in the product. Mixture experiments are a special type of response surface experiment that we will study in Chapter 11. They are very useful in designing biotechnology products, pharmaceuticals, foods and beverages, paints and coatings, consumer products such as detergents, soaps, and other personal care products, and a wide variety of other products.

### EXAMPLE 1.6 Designing a Web Page

A lot of business today is conducted via the World Wide Web. Consequently, the design of a business' web page has potentially important economic impact. Suppose that the Web site has the following components: (1) a photoflash image, (2) a main headline, (3) a subheadline, (4) a main text copy, (5) a main image on the right side, (6) a background design, and (7) a footer. We are interested in finding the factors that influence the click-through rate; that is, the number of visitors who click through into the site divided by the total number of visitors to the site. Proper selection of the important factors can lead to an optimal web page design. Suppose that there are four choices for the photoflash image, eight choices for the main headline, six choices for the subheadline, five choices for the main text copy,

four choices for the main image, three choices for the background design, and seven choices for the footer. If we use a factorial design, web pages for all possible combinations of these factor levels must be constructed and tested. This is a total of  $4 \times 8 \times 6 \times 5 \times 4 \times 3 \times 7 = 80,640$  web pages. Obviously, it is not feasible to design and test this many combinations of web pages, so a complete factorial experiment cannot be considered. However, a fractional factorial experiment that uses a small number of the possible web page designs would likely be successful. This experiment would require a fractional factorial where the factors have different numbers of levels. We will discuss how to construct these designs in Chapter 9.

## 1.3 Basic Principles

If an experiment such as the ones described in Examples 1.1 through 1.6 is to be performed most efficiently, a scientific approach to planning the experiment must be employed. **Statistical design of experiments** refers to the process of planning the experiment so that appropriate data will be collected and analyzed by statistical methods, resulting in valid and objective conclusions. The statistical approach to experimental design is necessary if we wish to draw meaningful conclusions from the data. When the problem involves data that are subject to experimental errors, statistical methods are the only **objective** approach to analysis. Thus, there are two aspects to any experimental problem: the design of the experiment and the statistical analysis of the data. These two subjects are closely related because the method

of analysis depends directly on the design employed. Both topics will be addressed in this book.

The three basic principles of experimental design are **randomization**, **replication**, and **blocking**. Sometimes we add the **factorial principle** to these three. Randomization is the cornerstone underlying the use of statistical methods in experimental design. By randomization we mean that both the allocation of the experimental material and the order in which the individual runs of the experiment are to be performed are randomly determined. Statistical methods require that the observations (or errors) be independently distributed random variables. Randomization usually makes this assumption valid. By properly randomizing the experiment, we also assist in "averaging out" the effects of extraneous factors that may be present. For example, suppose that the specimens in the hardness experiment are of slightly different thicknesses and that the effectiveness of the quenching medium may be affected by specimen thickness. If all the specimens subjected to the oil quench are thicker than those subjected to the saltwater quench, we may be introducing systematic bias into the experimental results. This bias handicaps one of the quenching media and consequently invalidates our results. Randomly assigning the specimens to the quenching media alleviates this problem.

Computer software programs are widely used to assist experimenters in selecting and constructing experimental designs. These programs often present the runs in the experimental design in random order. This random order is created by using a random number generator. Even with such a computer program, it is still often necessary to assign units of experimental material (such as the specimens in the hardness example mentioned above), operators, gauges or measurement devices, and so forth for use in the experiment.

Sometimes experimenters encounter situations where randomization of some aspect of the experiment is difficult. For example, in a chemical process, temperature may be a very hard-to-change variable as we may want to change it less often than we change the levels of other factors. In an experiment of this type, **complete randomization** would be difficult because it would add time and cost. There are statistical design methods for dealing with restrictions on randomization. Some of these approaches will be discussed in subsequent chapters (see in particular Chapter 14).

By **replication** we mean an independent **repeat run** of each factor combination. In the metallurgical experiment discussed in Section 1.1, replication would consist of treating a specimen by oil quenching and treating a specimen by saltwater quenching. Thus, if five specimens are treated in each quenching medium, we say that five **replicates** have been obtained. Each of the 10 observations should be run in random order. Replication has two important properties. First, it allows the experimenter to obtain an estimate of the experimental error. This estimate of error becomes a basic unit of measurement for determining whether observed differences in the data are really *statistically* different. Second, if the sample mean ( $\bar{y}$ ) is used to estimate the true mean response for one of the factor levels in the experiment, replication permits the experimenter to obtain a more precise estimate of this parameter. For example; if  $\sigma^2$  is the variance of an individual observation and there are  $n$  replicates, the variance of the sample mean is

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}$$

The practical implication of this is that if we had  $n = 1$  replicates and observed  $y_1 = 145$  (oil quench) and  $y_2 = 147$  (saltwater quench), we would probably be unable to make satisfactory inferences about the effect of the quenching medium—that is, the observed difference could be the result of experimental error. The point is that without replication we have no way of knowing why the two observations are different. On the other hand, if  $n$  was reasonably large and the experimental error was sufficiently small and if we observed sample averages  $\bar{y}_1 < \bar{y}_2$ , we would be reasonably safe in concluding that

saltwater quenching produces a higher hardness in this particular aluminum alloy than does oil quenching.

Often when the runs in an experiment are randomized, two (or more) consecutive runs will have exactly the same levels for some of the factors. For example, suppose we have three factors in an experiment: pressure, temperature, and time. When the experimental runs are randomized, we find the following:

Run number	Pressure (psi)	Temperature (°C)	Time (min)
$i$	30	100	30
$i + 1$	30	125	45
$i + 2$	40	125	45

Notice that between runs  $i$  and  $i + 1$ , the levels of pressure are identical and between run  $i + 1$  and  $i + 2$ , the levels of both temperature and time are identical. To obtain a true replicate, the experimenter needs to “twist the pressure knob” to an intermediate setting between runs  $i$  and  $i + 1$ , and reset pressure to 30 psi for run  $i + 1$ . Similarly, temperature and time should be reset to intermediate levels between runs  $i + 1$  and  $i + 2$  before being set to their design levels for run  $i + 2$ . Part of the experimental error is the variability associated with hitting and holding factor levels.

There is an important distinction between **replication** and **repeated measurements**. For example, suppose that a silicon wafer is etched in a single-wafer plasma etching process and a critical dimension (CD) on this wafer is measured three times. These measurements are not replicates; they are a form of repeated measurements, and in this case the observed variability in the three repeated measurements is a direct reflection of the inherent variability in the measurement system or gauge and possibly the variability in this CD at different locations on the wafer where the measurement were taken. As another illustration, suppose that as part of an experiment in semiconductor manufacturing four wafers are processed simultaneously in an oxidation furnace at a particular gas flow rate and time and then a measurement is taken on the oxide thickness of each wafer. Once again, the measurements on the four wafers are not replicates but repeated measurements. In this case, they reflect differences among the wafers and other sources of variability within that particular furnace run. Replication reflects sources of variability both **between runs** and (potentially) **within runs**.

**Blocking** is a design technique used to improve the precision with which comparisons among the factors of interest are made. Often blocking is used to reduce or eliminate the variability transmitted from **nuisance factors**—that is, factors that may influence the experimental response but in which we are not directly interested. For example, an experiment in a chemical process may require two batches of raw material to make all the required runs. However, there could be differences between the batches due to supplier-to-supplier variability, and if we are not specifically interested in this effect, we would think of the batches of raw material as a nuisance factor. Generally, a block is a set of relatively homogeneous experimental conditions. In the chemical process example, each batch of raw material would form a block, because the variability within a batch would be expected to be smaller than the variability between batches. Typically, as in this example, each level of the nuisance factor becomes a block. Then the experimenter divides the observations from the statistical design into groups that are run in each block. We study blocking in detail in several places in the text, including Chapters 4, 5, 7, 8, 9, 11, and 13. A simple example illustrating the blocking principle is given in Section 2.5.1.

The three basic principles of experimental design, randomization, replication, and blocking are part of every experiment. We will illustrate and emphasize them repeatedly throughout this book.