1

The Role of Statistics in Engineering

CHAPTER OUTLINE

LEARNING OBJECTIVES

After careful study of this chapter you should be able to do the following:

1. Identify the role that statistics can play in the engineering problem-solving process
2. Discuss how variability affects the data collected and used for making engineering decisions
3. Explain the difference between enumerative and analytical studies
4. Discuss the different methods that engineers use to collect data
5. Identify the advantages that designed experiments have in comparison to other methods of collecting engineering data
6. Explain the differences between mechanistic models and empirical models
7. Discuss how probability and probability models are used in engineering and science

1-1 The Engineering Method and Statistical Thinking

An engineer is someone who solves problems of interest to society by the efficient application of scientific principles by

- Refining existing products
- Designing new products or processes

1-1 The Engineering Method and Statistical Thinking

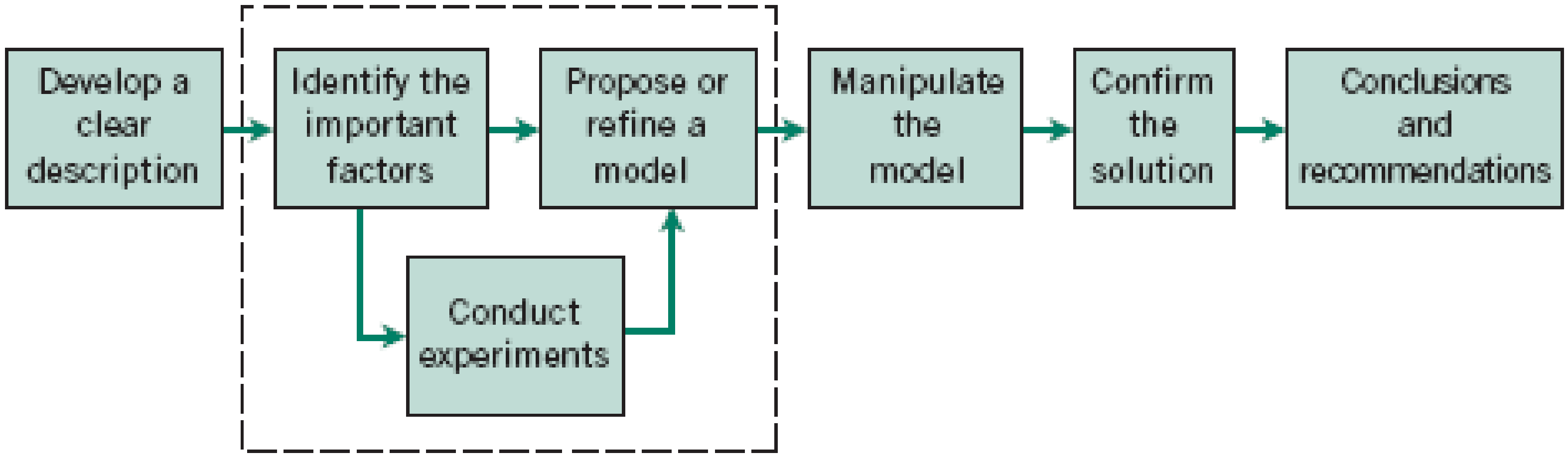

Figure 1.1 The engineering method

1-1 The Engineering Method and Statistical Thinking

The field of statistics deals with the collection, presentation, analysis, and use of data to

- Make decisions
- Solve problems
- Design products and processes

1-1 The Engineering Method and Statistical Thinking

- Statistical techniques are useful for describing and understanding variability.
- By variability, we mean successive observations of a system or phenomenon do *not* produce exactly the same result.
- Statistics gives us a framework for describing this variability and for learning about potential sources of variability.

1-1 The Engineering Method and Statistical Thinking

Engineering Example

An engineer is designing a nylon connector to be used in an automotive engine application. The engineer is considering establishing the design specification on wall thickness at 3/32 inch but is somewhat uncertain about the effect of this decision on the connector pull-off force. If the pull-off force is too low, the connector may fail when it is installed in an engine. Eight prototype units are produced and their pull-off forces measured (in pounds): 12.6, 12.9, 13.4, 12.3, 13.6, 13.5, 12.6, 13.1.

1-1 The Engineering Method and Statistical Thinking

Engineering Example

- The **dot diagram** is a very useful plot for displaying a small body of data - say up to about 20 observations.
- This plot allows us to see easily two features of the data; the **location**, or the middle, and the **scatter** or **variability**.

Figure 1-2 Dot diagram of the pull-off force data when wall thickness is 3/32 inch.

1-1 The Engineering Method and Statistical Thinking

Engineering Example

• The engineer considers an alternate design and eight prototypes are built and pull-off force measured.
• The dot diagram can be used to compare two sets of data

Figure 1-3 Dot diagram of pull-off force for two wall thicknesses.

1-1 The Engineering Method and Statistical Thinking

Engineering Example

- Since pull-off force varies or exhibits variability, it is a random variable.

- A random variable, X, can be model by

$$X = \mu + \varepsilon$$

where μ is a constant and ε a random disturbance.

1-1 The Engineering Method and Statistical Thinking

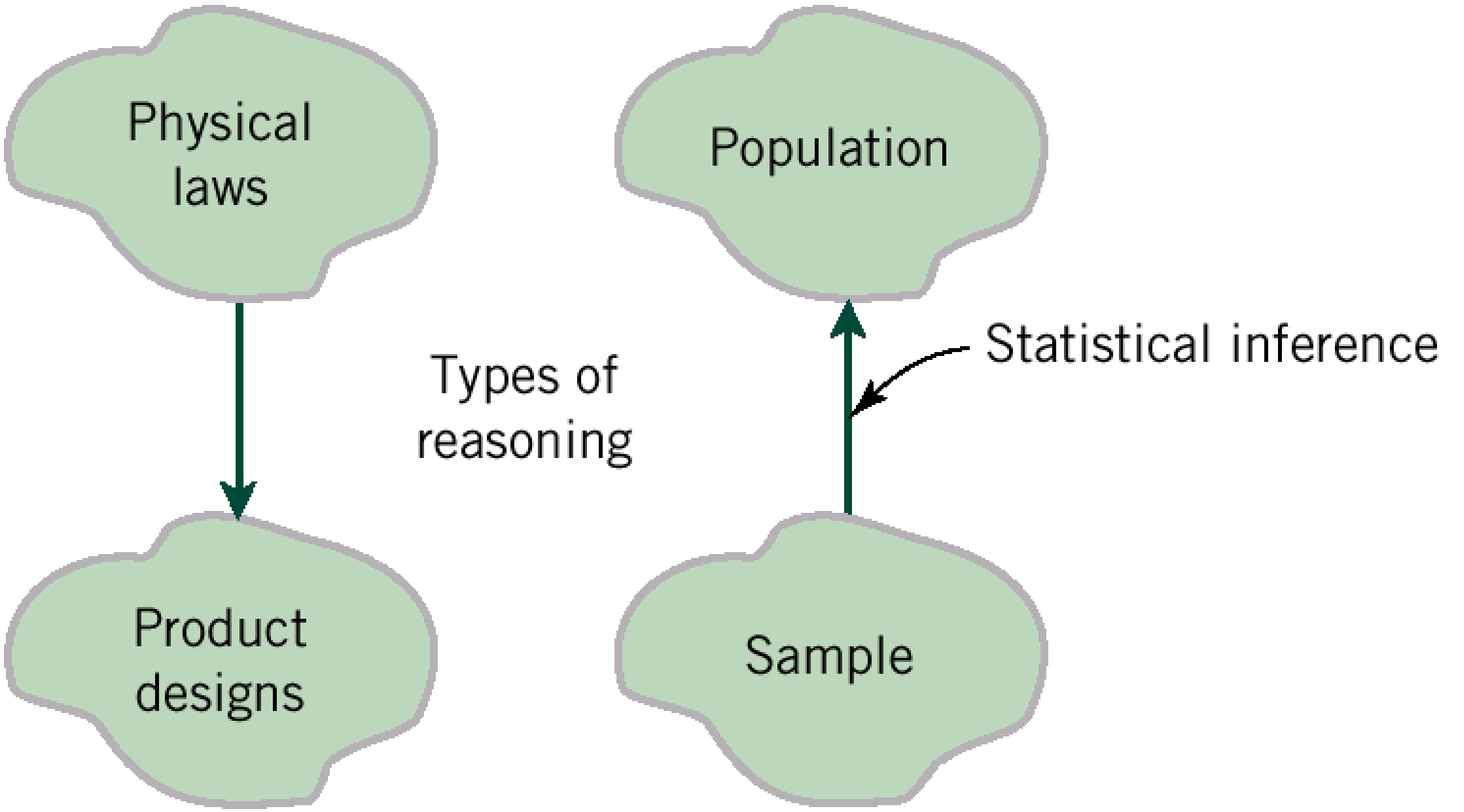

Figure 1-4 Statistical inference is one type of reasoning.

1-2 Collecting Engineering Data

Three basic methods for collecting data:

- A **retrospective** study using historical data
- An **observational** study
- A **designed experiment**

1-2.4 Designed Experiments

Table 1-1 The Designed Experiment (Factorial Design) for the Distillation Column

Reboil Temp.	Condensate Temp.	Reflux Rate
−1	−1	−1
+1	−1	−1
−1	+1	−1
+1	+1	−1
−1	−1	+1
+1	−1	+1
−1	+1	+1
+1	+1	+1

1-2.4 Designed Experiments

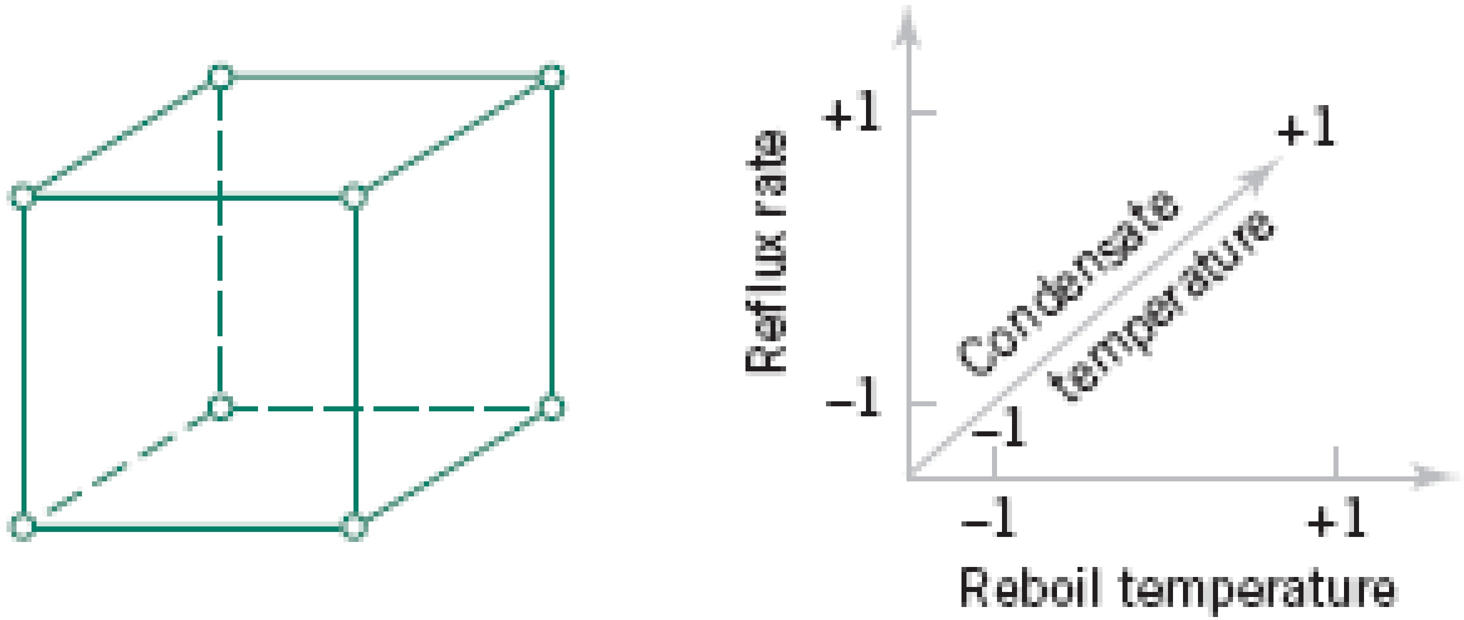

Figure 1-5 The factorial design for the distillation column

1-2.4 Designed Experiments

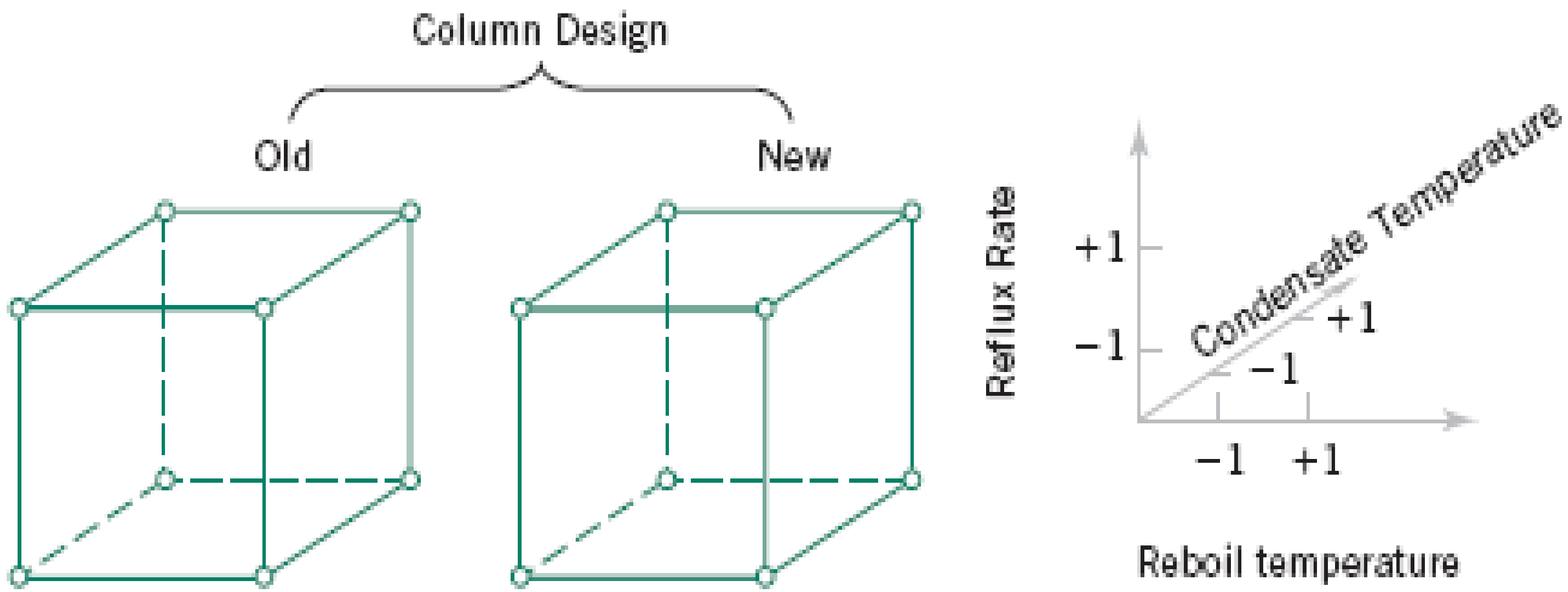

Figure 1-6 A four-factorial experiment for the distillation column

1-2.5 Observing Processes Over Time

Whenever data are collected over time it is important to plot the data over time. Phenomena that might affect the system or process often become more visible in a time-oriented plot and the concept of stability can be better judged.

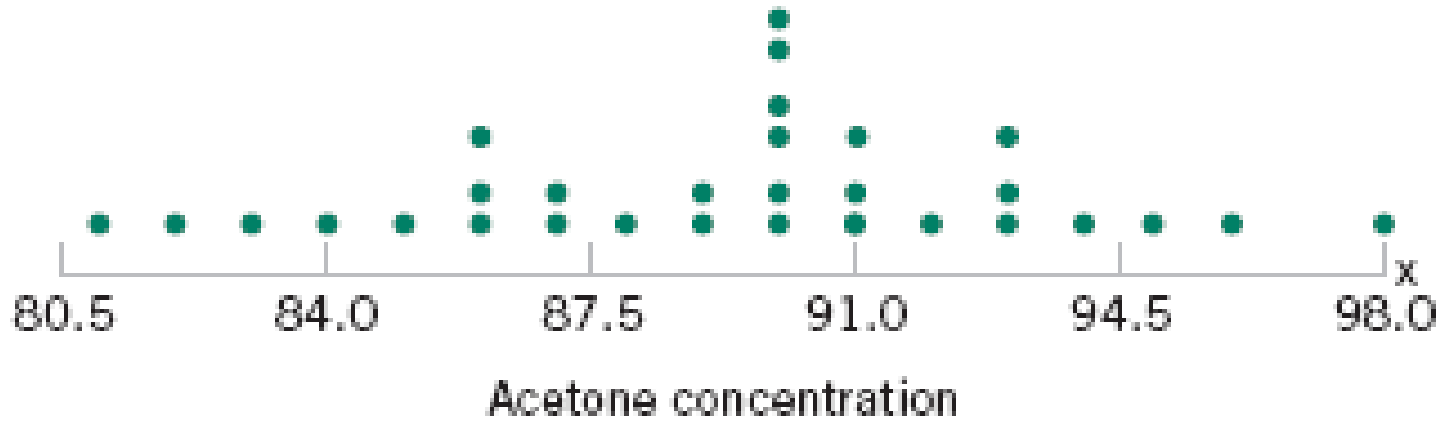

Figure 1-8 The dot diagram illustrates variation but does not identify the problem.

1-2.5 Observing Processes Over Time

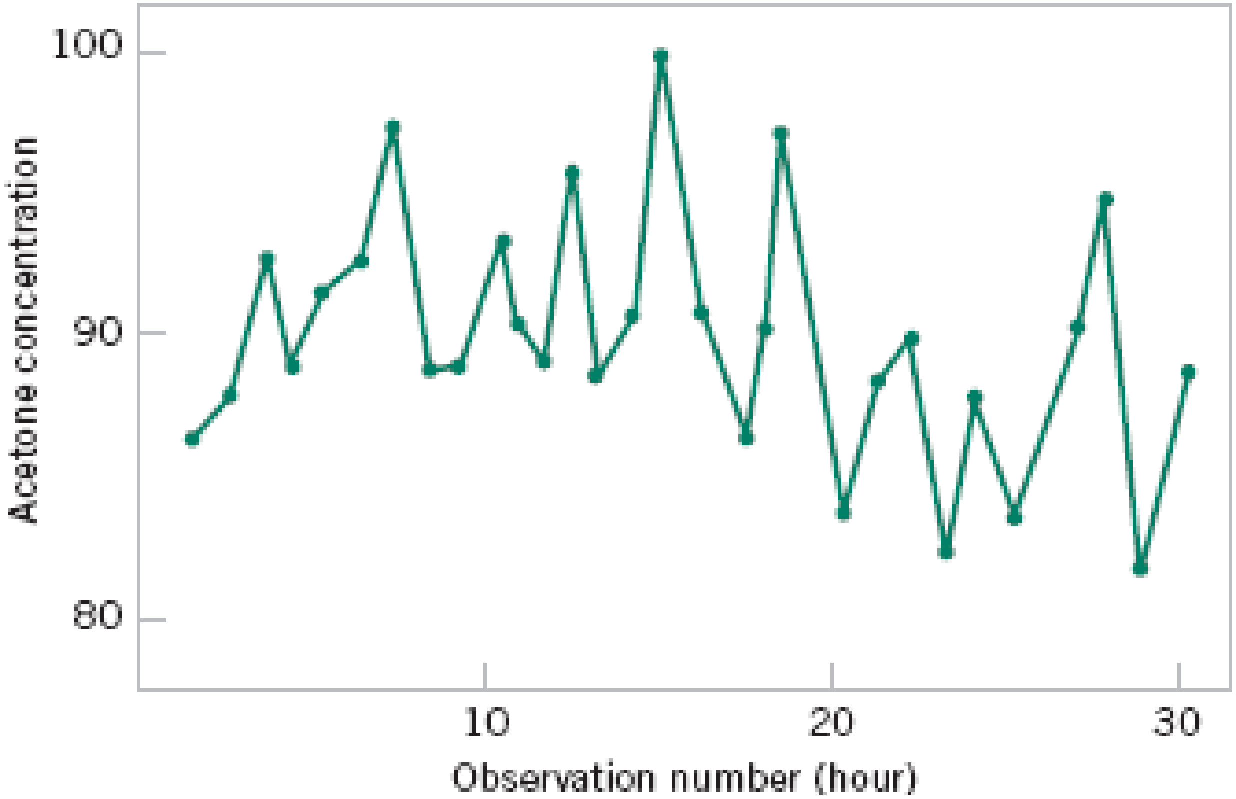

Figure 1-9 A time series plot of concentration provides more information than a dot diagram.

1-2.5 Observing Processes Over Time

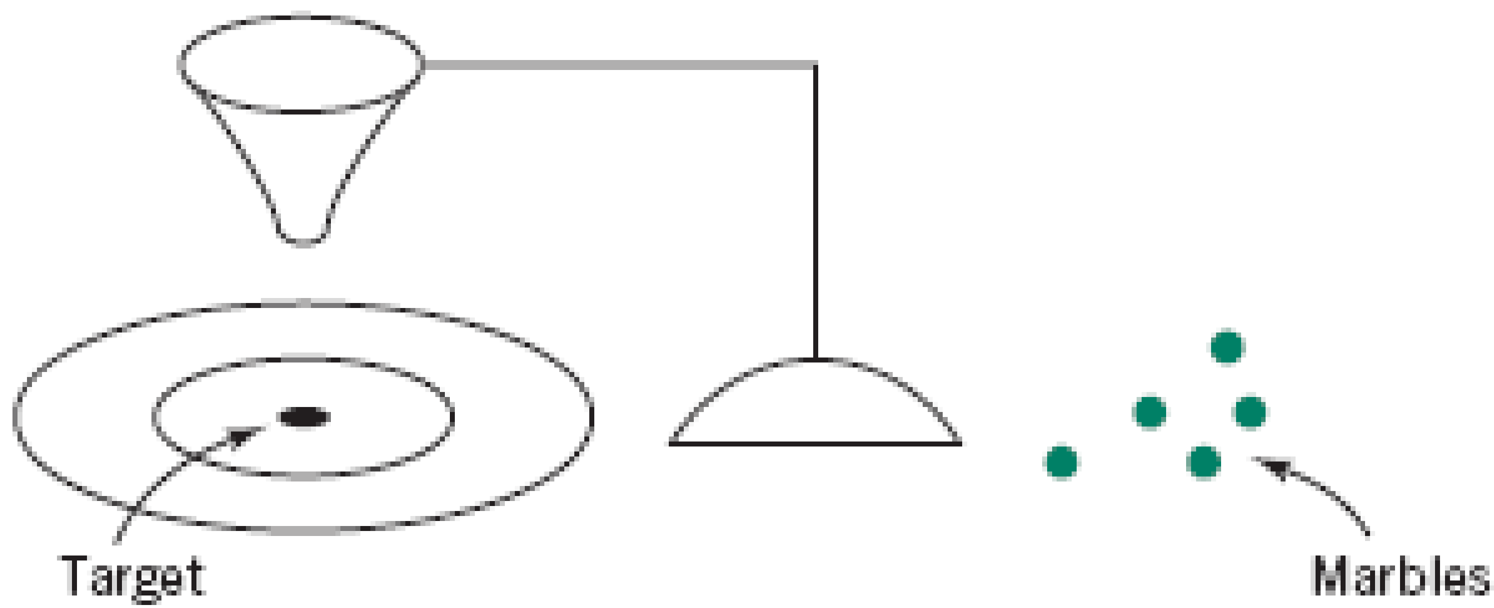

Figure 1-10 Deming's funnel experiment.

1-2.5 Observing Processes Over Time

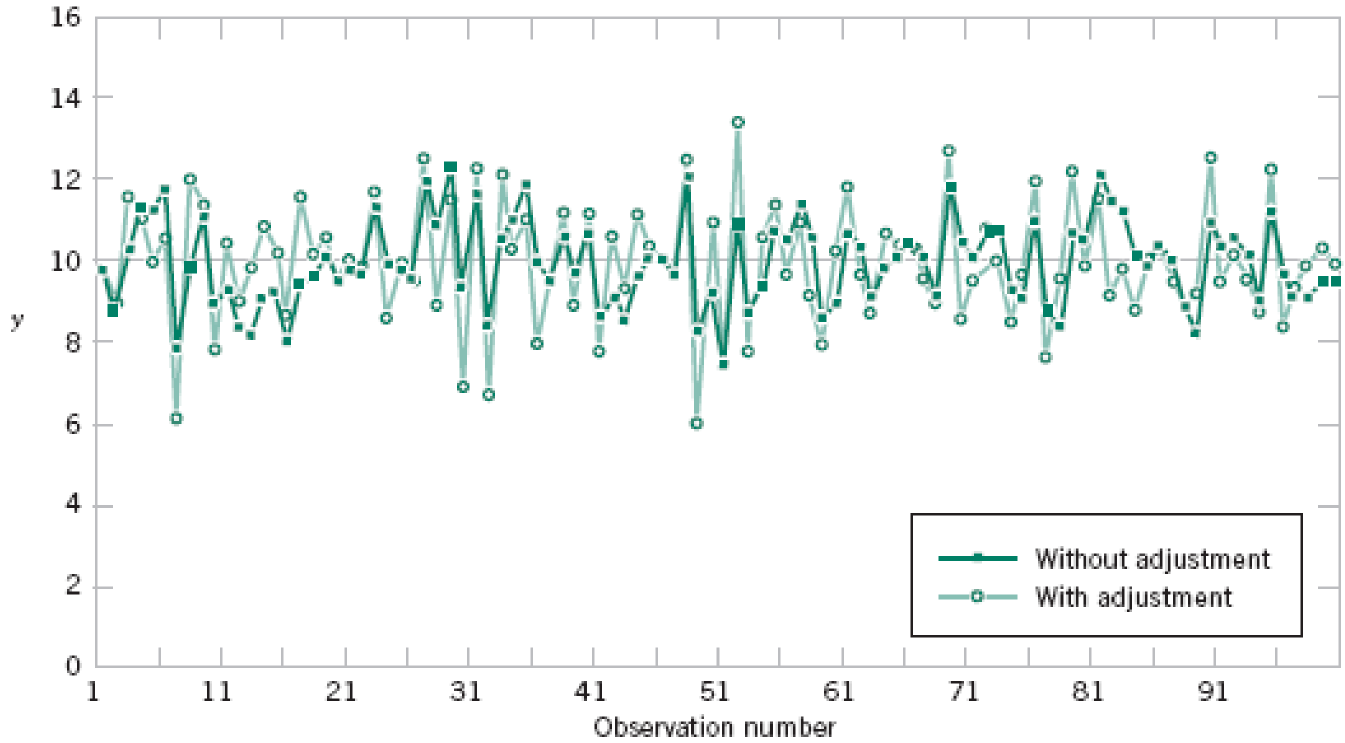

Figure 1-11 Adjustments applied to random disturbances over control the process and increase the deviations from the target.

1-2.5 Observing Processes Over Time

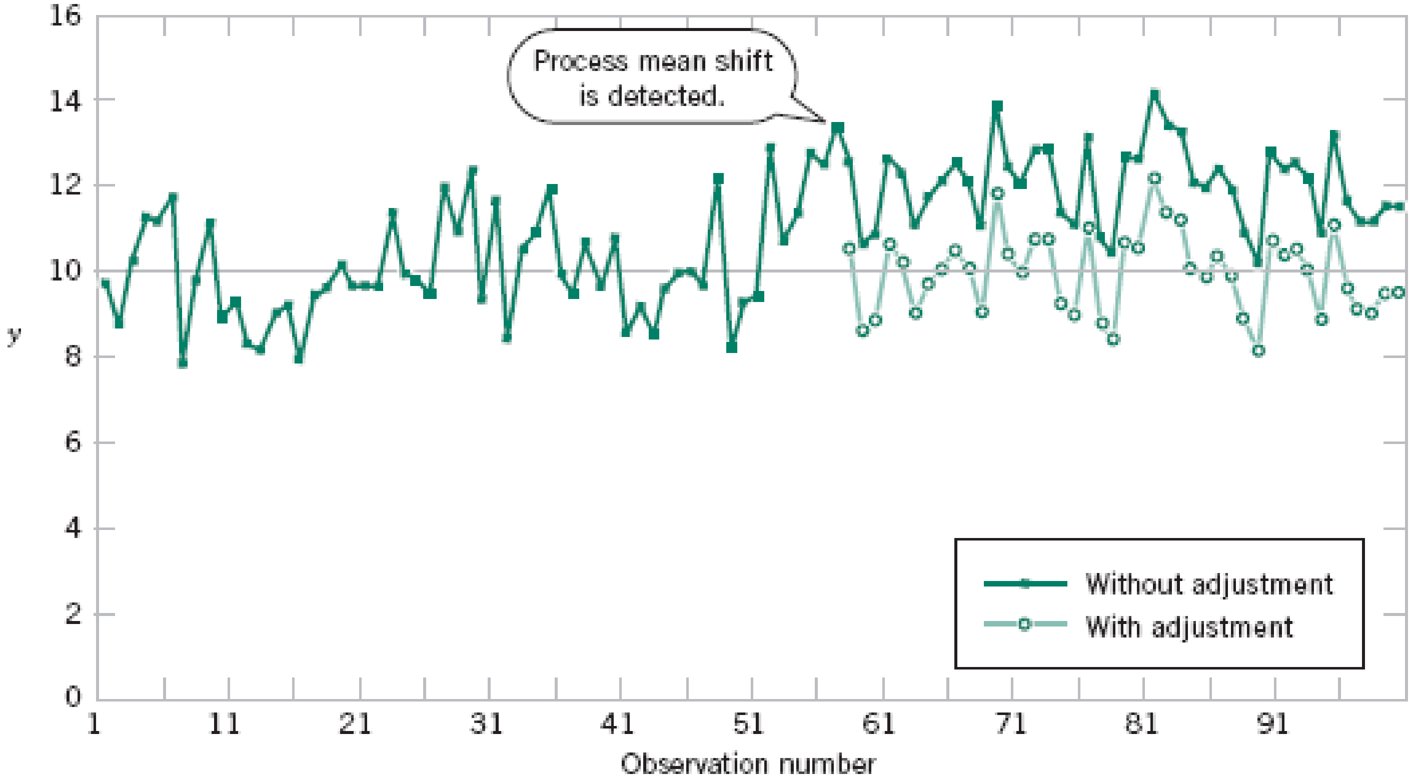

Figure 1-12 Process mean shift is detected at observation number 57, and one adjustment (a decrease of two units) reduces the deviations from target.

1-2.6 Observing Processes Over Time

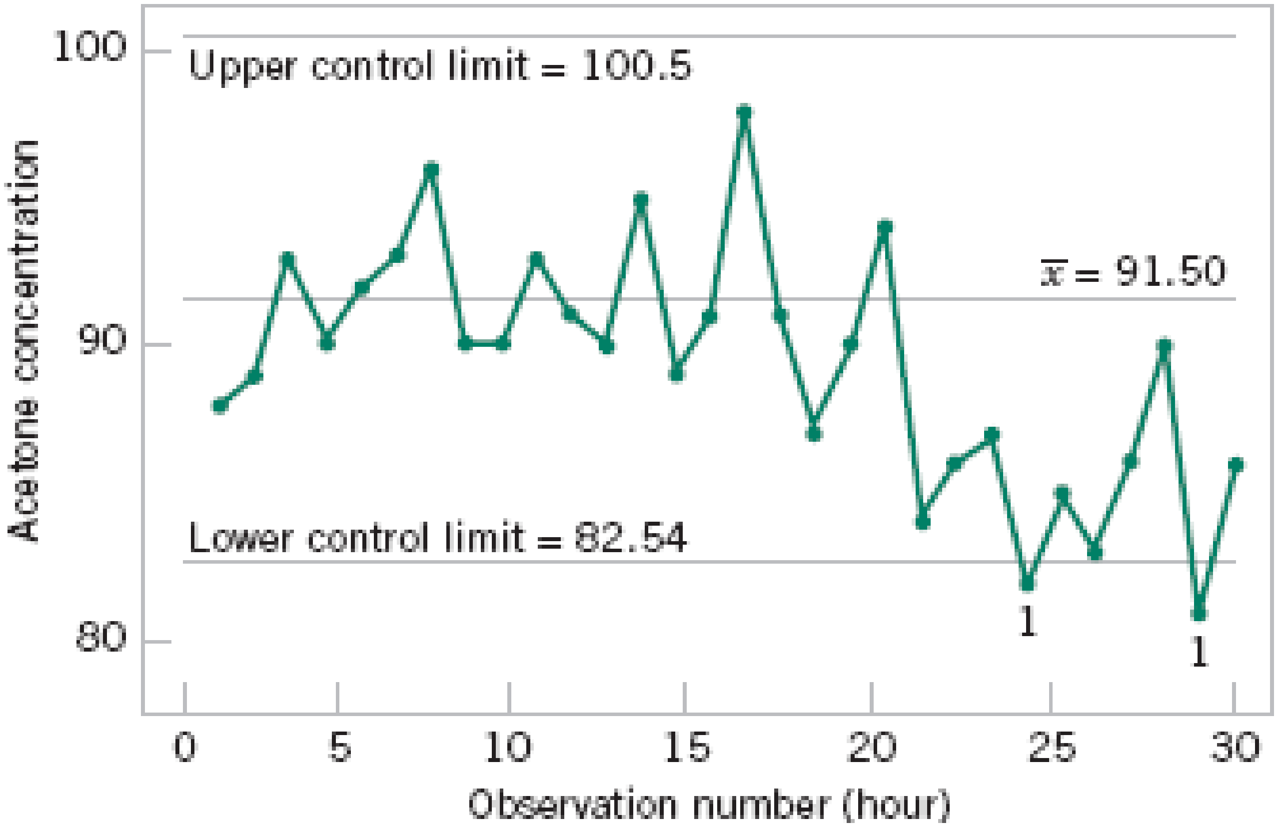

Figure 1-13 A control chart for the chemical process concentration data.

1-2.6 Observing Processes Over Time

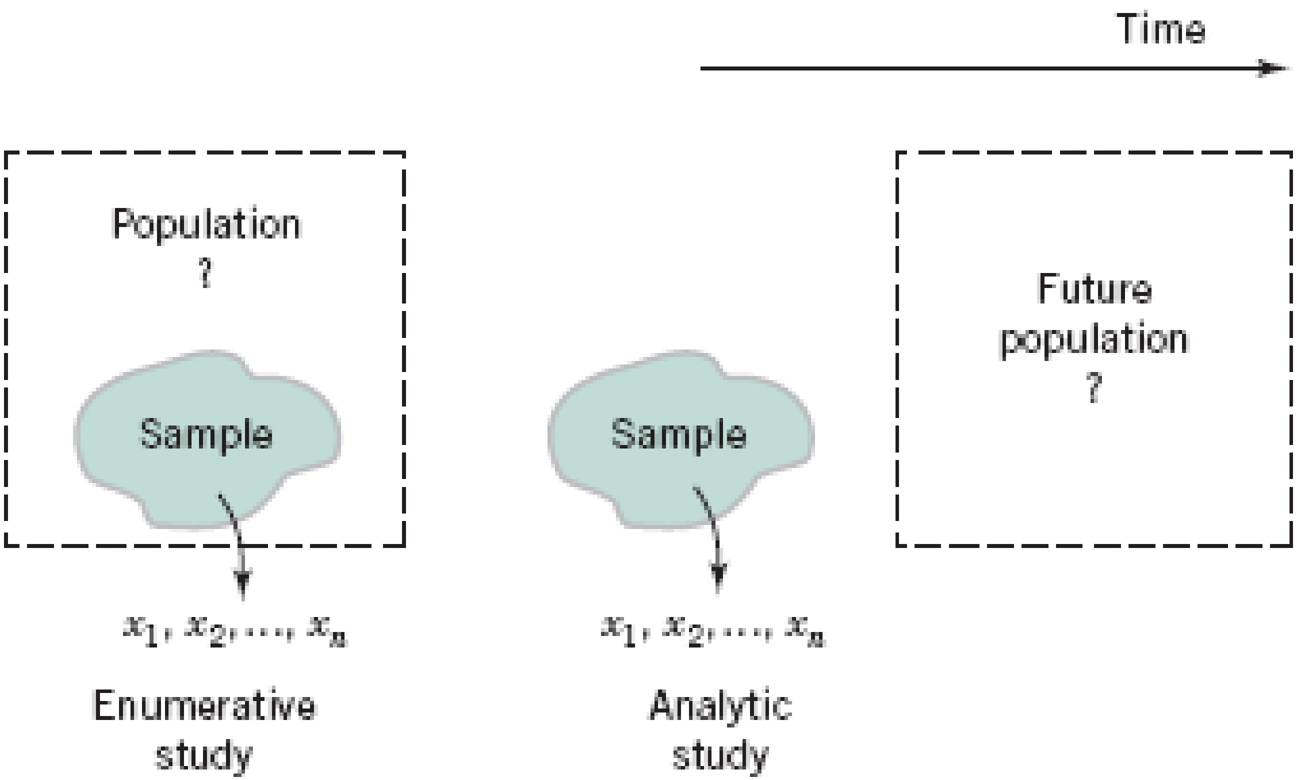

Figure 1-14 Enumerative versus analytic study.

1-3 Mechanistic and Empirical Models

A **mechanistic model** is built from our underlying knowledge of the basic physical mechanism that relates several variables.

Example: Ohm's Law

Current = voltage/resistance

$$I = E/R$$

$$I = E/R + \varepsilon$$

1-3 Mechanistic and Empirical Models

An **empirical model** is built from our engineering and scientific knowledge of the phenomenon, but is not directly developed from our theoretical or first-principles understanding of the underlying mechanism.

1-3 Mechanistic and Empirical Models

Example

Suppose we are interested in the number average molecular weight (M_n) of a polymer. Now we know that M_n is related to the viscosity of the material (V), and it also depends on the amount of catalyst (C) and the temperature (T) in the polymerization reactor when the material is manufactured. The relationship between M_n and these variables is $M_n = f(V, C, T)$

say, where the *form* of the function f is unknown.

$$M_n = \beta_0 + \beta_1 V + \beta_2 C + \beta_3 T + \epsilon$$

where the β's are unknown parameters.

Table 1-2 Wire Bond Pull Strength Data

Observation Number	Pull Strength y	Wire Length x_1	Die Height x_2
1	9.95	2	50
2	24.45	8	110
3	31.75	11	120
4	35.00	10	550
5	25.02	8	295
6	16.86	4	200
7	14.38	2	375
8	9.60	2	52
9	24.35	9	100
10	27.50	8	300
11	17.08	4	412
12	37.00	11	400
13	41.95	12	500
14	11.66	2	360
15	21.65	4	205
16	17.89	4	400
17	69.00	20	600
18	10.30	1	585
19	34.93	10	540
20	46.59	15	250
21	44.88	15	290
22	54.12	16	510
23	56.63	17	590
24	22.13	6	100
25	21.15	5	400

1-3 Mechanistic and Empirical Models

$$\text{Pull strength} = \beta_0 + \beta_1(\text{wire length}) + \beta_2(\text{die height}) + \epsilon$$

In general, this type of empirical model is called a **regression model**.

The **estimated** regression line is given by

$$\widehat{\text{Pull strength}} = 2.26 + 2.74(\text{wire length}) + 0.0125(\text{die height})$$

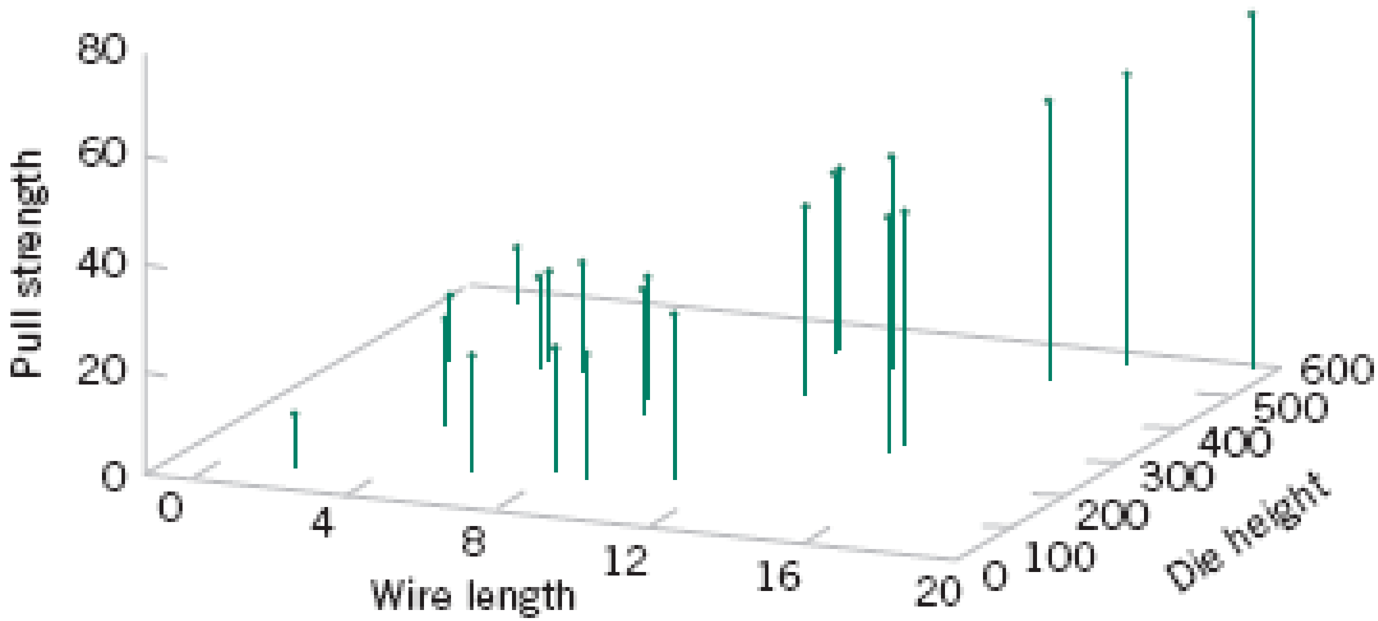

Figure 1-15 Three-dimensional plot of the wire and pull strength data.

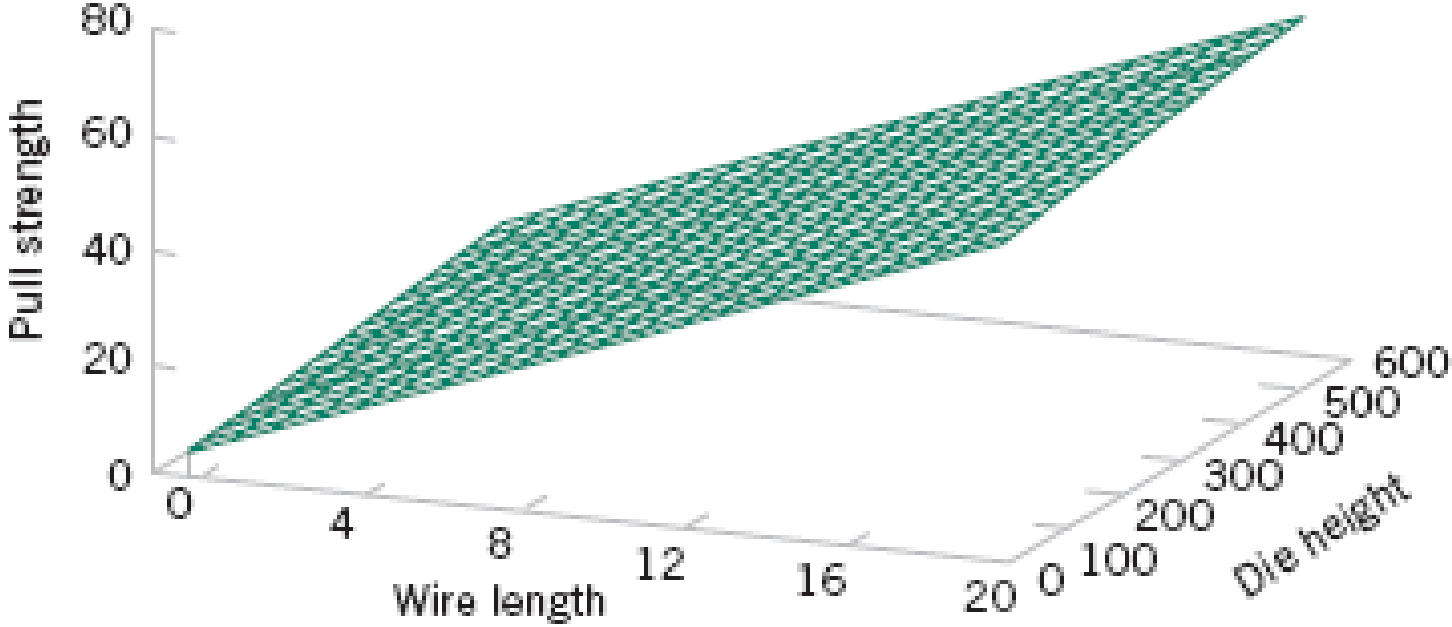

Figure 1-16 Plot of the predicted values of pull strength from the empirical model.

1-4 Probability and Probability Models

- **Probability models** help quantify the risks involved in statistical inference, that is, risks involved in decisions made every day.
- Probability provides the **framework** for the study and application of statistics.

IMPORTANT TERMS AND CONCEPTS

- Analytic study
- Cause and effect
- Designed experiment
- Empirical model
- Engineering method
- Enumerative study
- Factorial Experiment
- Fractional factorial experiment
- Hypothesis testing
- Interaction
- Mechanistic model
- Observational study
- Overcontrol
- Population
- Probability model
- Problem-solving method
- Randomization
- Retrospective study
- Sample
- Statistical inference
- Statistical Process Control
- Statistical thinking
- Tampering
- Time series
- Variability