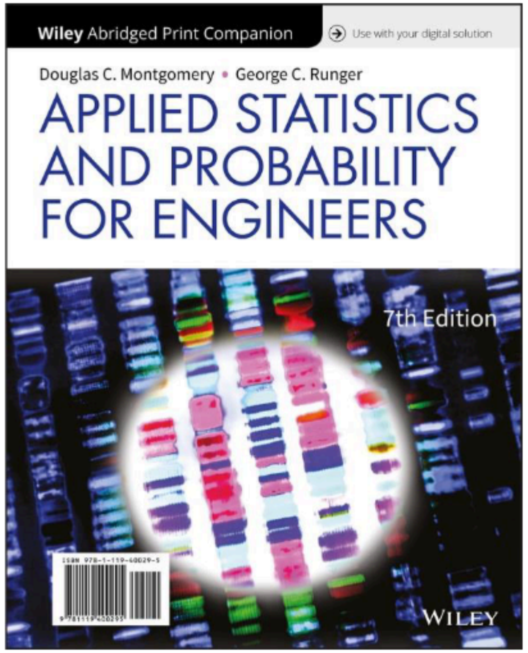
WILEY



Applied Statistics and Probability for Engineers

Seventh Edition

Douglas C. Montgomery   George C. Runger

Chapter 1

The Role of Statistics in Engineering

Chapter 1 Title Slide

1 The Role of Statistics in Engineering

CHAPTER OUTLINE

WILEY

Learning Objectives for Chapter 1

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 the other methods of collecting engineering data
6. Explain the differences between mechanistic models & empirical models
7. Discuss how probability and probability models are used in engineering and science

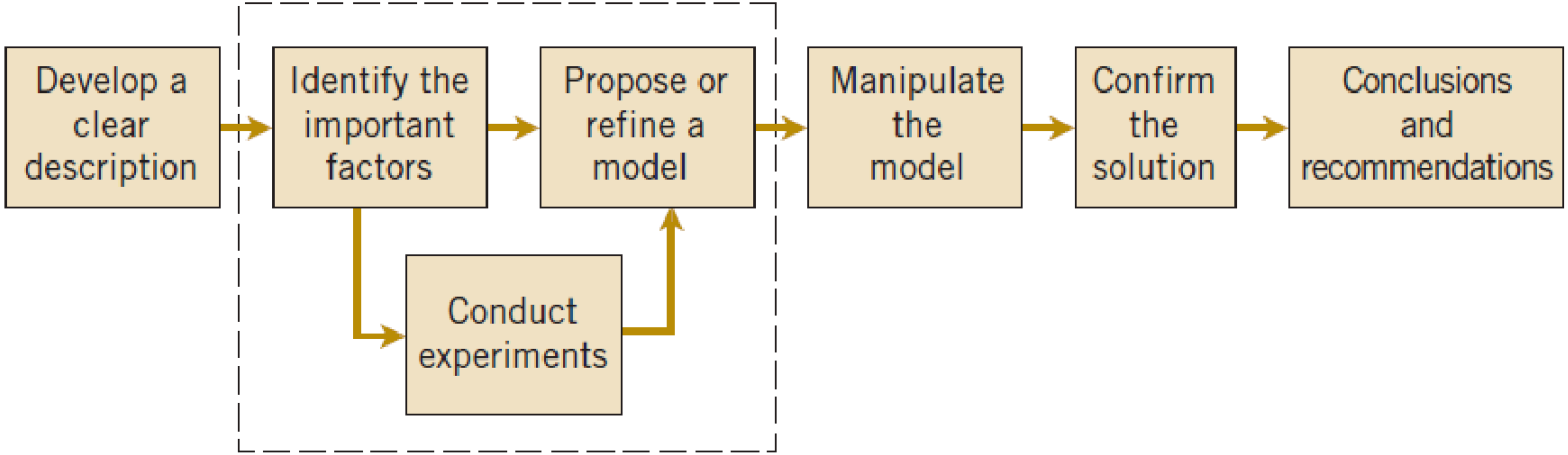
What Do Engineers Do?

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

- Refining existing product or process
- Designing new product or process

that meets customers' needs

The Engineering Method



WILEY

Statistics Supports the Engineering Method

- The field of **statistics** deals with the collection, presentation, analysis, and use of data to:
  - Make decisions
  - Solve problems
  - Design products and processes
- It is the **science of data**.

WILEY

Variability

- Statistical methods are useful to describe and understand **variability**.
- By variability, we mean successive observations of a system or phenomenon do *not* produce exactly the same result.
- Incorporating variability into decision-making processes is called **statistical thinking.**

Statistics provides a framework for describing this variability and for learning about potential **sources of variability**.

WILEY

An Engineering Example of Variability

- Eight prototype units are produced and their pull-off forces are measured (in pounds): $12.6,\ 12.9,\ 13.4,\ 12.3,\ 13.6,\ 13.5,\ 12.6,\ 13.1$
- All of the prototypes do not have the same pull-off force. We can see the variability in the above measurements as they exhibit variability.
- 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**.

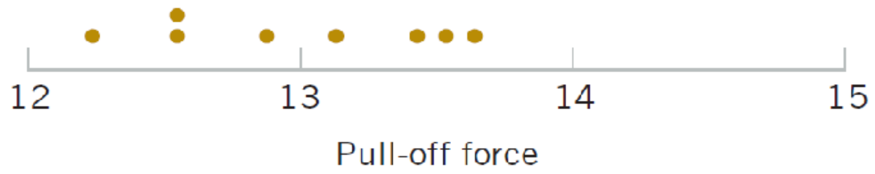


Figure 1.2 Dot diagram of the pull-off force data when wall thickness is $3/32$ inch

WILEY

Basic Methods of Collecting Data

Three basic methods of collecting data:

- A **retrospective** study using historical data
  Data collected in the past for other purposes.
- An **observational** study
  Data, presently collected, by a passive observer.
- A **designed experiment**
  Data collected in response to process input changes.

WILEY

Hypothesis Tests

- **Hypothesis Test**
  - A statement about some aspect of the system, e.g. process behavior value.
  - Compared to a claim about another process value.
  - Data are gathered to support or refute the claim.

- **One-sample hypothesis test**
  - Example: $mean\ strength = 12.75\ \ \boldsymbol{vs}.\ mean\ strength > 12.75$

- **Two-sample hypothesis test**
  - Example: $\frac{1}{8} in.\ mean\ strength - \frac{3}{32} in.\ mean\ strength = 0$

$$\boldsymbol{vs}.$$

$$\frac{1}{8} in.\ mean\ strength - \frac{3}{32} in.\ mean\ strength > 0$$

Factorial Experiment & Example

- An experimental design which uses every possible combination of the factor levels to form a basic experiment with k different settings for the process.
- **Factor levels** are specified values of factors used in the experiment
- This type of experiment is called a **factorial experiment.**
- **Example.** Consider a petroleum distillation column:
    - **Output** is acetone concentration
    - **Inputs** (factors) are:
        1. Reboil temperature
        2. Condensate temperature
        3. Reflux rate
    - Output changes as the inputs are changed by experimenter.

WILEY

Factorial Experiment Example

- Each factor is set at 2 reasonable levels (-1 and $+1$)
- 8 (2^3) runs are made, at every combination of factors, to observe acetone output.
- Resultant data is used to create a mathematical model of the process representing cause and effect.

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

Fractional Factorial Experiment

- Factorial experiments can get too large. For example, 8 factors will require $2^8 = 256$ experimental runs of the distillation column.
- Certain combinations of factor levels can be deleted from the experiments without degrading the resultant model.
- The result is called a **fractional factorial experiment**.

WILEY

Fractional Factorial Experiment Example

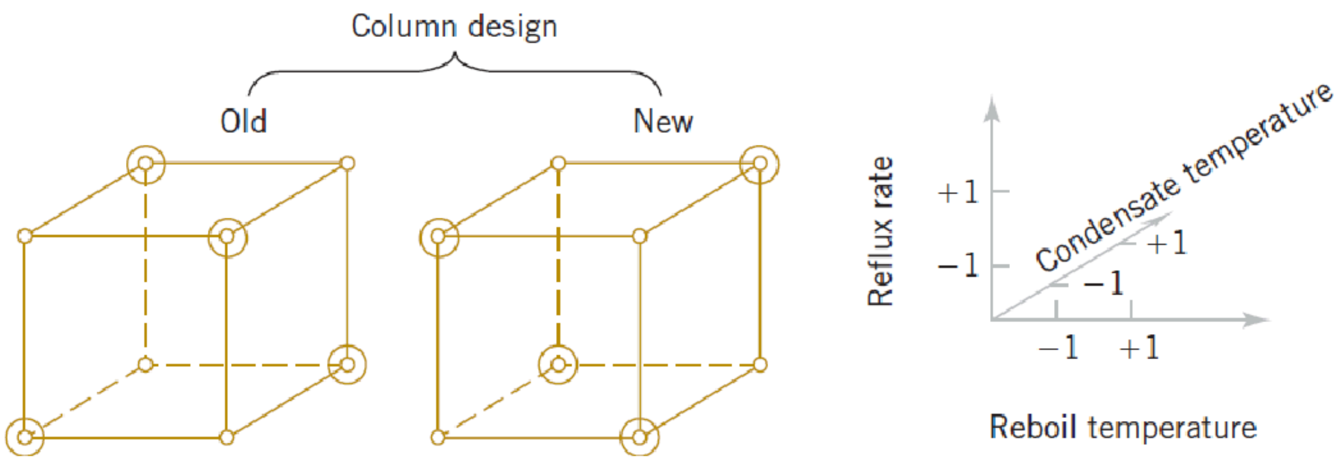


Figure 1.7 A fractional factorial experiment for the distillation column (one-half fraction) $2^4 / 2 = 8$ circled settings.

An Experiment in Variation

W. Edwards Deming, a famous industrial statistician & contributor to the Japanese quality revolution, conducted an illustrative experiment on process **over-control** or **tampering**.

Let's look at his apparatus and experimental procedure.

Deming's Experimental Set-up

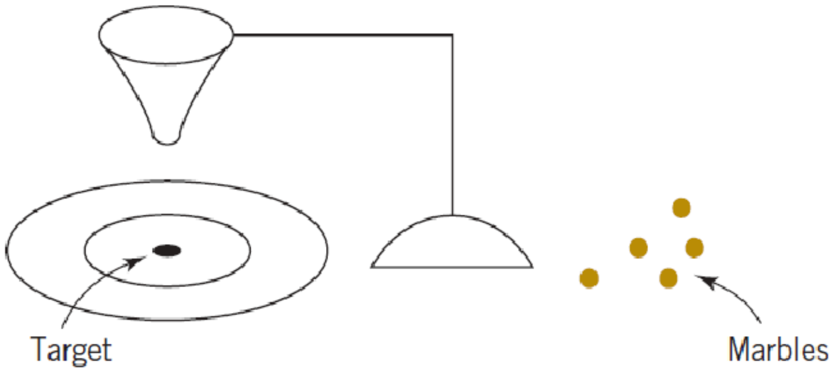


Figure 1.10 Deming's Funnel experiment

- Marbles were dropped through a funnel onto a target and the location where the marble struck the target was recorded.
- Variation was caused by several factors:
  - Marble placement in funnel & release dynamics
  - Vibration
  - Air currents
  - Measurement errors

WILEY

Deming's Experimental Procedure

The **funnel** was aligned with the center of the target. Marbles were dropped. The distance from the strike point to the target center was measured and recorded.

Strategy 1: Never move funnel. Repeat process.

Strategy 2: Move funnel at an equal distance in the opposite direction to compensate for the error. Continue to make this type of adjustment after each marble is dropped. Repeat process.

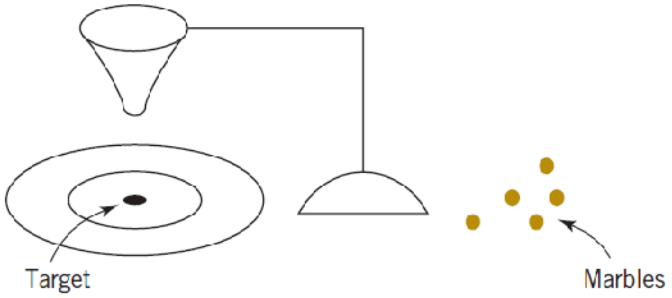


Figure 1.10 Deming's Funnel experiment

WILEY

Deming's Experimental Procedure

- When both strategies were completed, the variability of the distance from the target for strategy 2 was approximately twice as large than for strategy 1.
- The deviations from the target is increased due to the adjustments to the funnel.
  - Adjustments to the funnel do not decrease future errors.
  - Instead, they tend to move the funnel farther from the target.
- This experiment explains that the adjustments to a process based on random disturbances can actually increase the variation of the process. This is referred to as **overcontrol** or **tampering.**

WILEY

Conclusions from the Deming Experiment

- The lesson of the Deming experiment is that a process should not be adjusted in response to random variation, but only when a clear shift in the process value becomes apparent.

- Then a process adjustment should be made to return the process outputs to their normal values.

- To identify when the shift occurs, a **control chart** is used. Output values, plotted over time along with the outer limits of normal variation, pinpoint when the process leaves normal values and should be adjusted.

WILEY

How Is the Change Detected?

- A **control chart** is used. Its characteristics are:
  - Time-oriented horizontal axis, e.g., hours
  - Variable-of-interest vertical axis, e.g., acetone concentration
- Long-term average is plotted as the **center line**.
- Long-term usual variability is plotted as **upper** and **lower control limits** around the long-term average.
  - Limits typically located at 3 standard deviations away from center line
- A sample of size n is taken and the averages are plotted over time. If the plotted points are between the control limits, then the process is normal; if not, it needs to be adjusted.

WILEY

How Is the Change Detected Graphically?

The center line on the control chart is just the average of the concentration measurements for the first 20 samples

$$\bar{x} = 91.5\ g/l$$

when the process is stable. The upper control limit and the lower control limit are located 3 standard deviations of the concentration values above and below the center line.

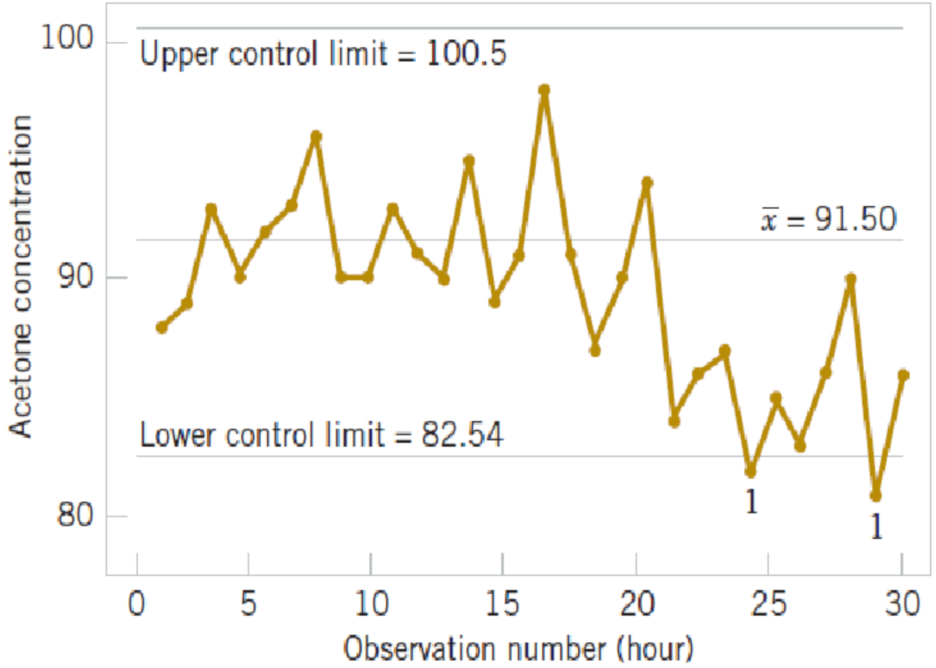


Figure 1.13 A control chart for the chemical process concentration data. Process goes out of limits at hours 24 & 29. Shut down & adjust process.

WILEY

Use of Control Charts

Deming contrasted two purposes of control charts:

1. **Enumerative studies**: Control chart of past production lots. Used for lot-by-lot acceptance sampling.

2. **Analytic studies**: Real-time control of a production process.

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***

$$\begin{aligned} Current &= Voltage/Resistance \\ I &= E/R \\ I &= E/R + \epsilon \end{aligned}$$

where ϵ is a term added to the model to account for the fact that the observed values of current flow do not perfectly conform to the mechanistic model

The form of the function is **known**.

WILEY

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.

The form of the function is **not known** *a priori*.

WILEY

An Example of an Empirical Model

- In a semiconductor manufacturing plant, the finished semiconductor is wire-bonded to a frame. In an observational study, the variables recorded were:
    - Pull strength to break the bond (y)
    - Wire length (x_1)
    - Die height (x_2)
- The data recorded are shown on the next slide.

WILEY

An Example of an Empirical Model

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

An Example of an Empirical Model

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

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

- The **estimated** regression relationship is given by:

$$Pull\ s\widehat{treng}th = 2.26 + 2.74(wire\ length) + 0.0125(die\ height)$$

where the "hat," or circumflex, over pull strength indicates that this is an estimated or predicted quality

WILEY

Visualizing the Data

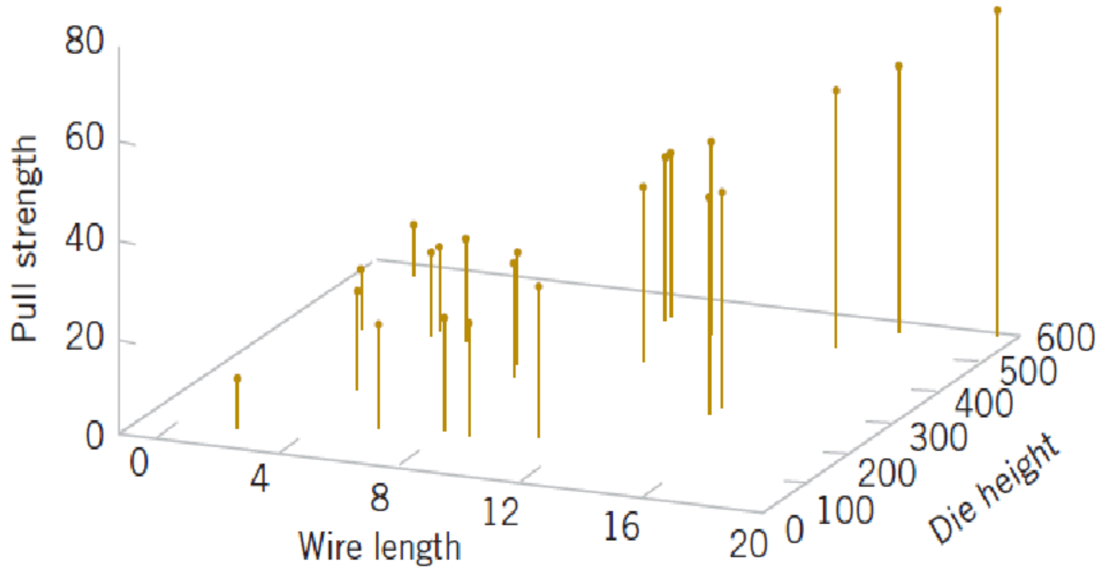


Figure 1.15 Three-dimensional plot of the pull strength (y), wire length (x_1) and die height (x_2) data.

Visualizing the Resultant Model Using Regression Analysis

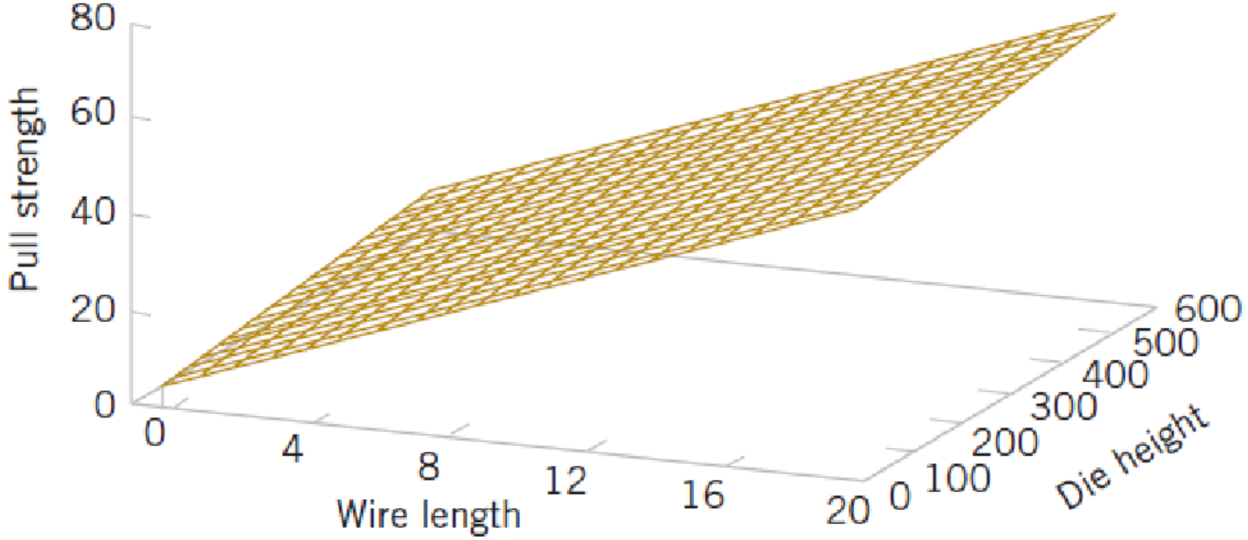


Figure 1.16 Plot of the predicted values (a plane) of pull strength from the empirical regression model

WILEY

Models Can Also Reflect Uncertainty

- **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.
- Probability concepts will be introduced in the next lecture.

WILEY

Important Terms & Concepts of Chapter 1

Analytic study
Cause and effect
Designed experiment
Empirical model
Engineering method
Enumerative study
Factorial experiment
Fractional factorial experiment
Hypothesis
Hypothesis testing
Interaction
Mechanistic model
Observational study
Overcontrol
Population
Probability model
Random variable
Randomization
Retrospective study
Sample
Scientific method
Statistical inference
Statistical process control
Statistical thinking
Tampering
Time series
Variability