Chapter 1  Introduction
1.1 Engineering Statistics

- Planning for and collecting data
- Summarizing data
- Drawing conclusions based on data
1.2 Terminology

1.2.1 Types of Statistical Studies

- **Observational study** – We just observe the process
- **Experimental study** – We purposefully manipulate the process

Compression strengths of taconite pellets
- Observe historical effects of changing kiln temperature
- Purposefully run at set kiln temperatures
Firmer conclusions from an experimental study. Other **lurking variables** can obscure effects in an observational study.

- Kiln temperature may have been adjusted to compensate for ambient temperature or other lurking variables
- Harder to isolate effects of kiln temperature if other variables are changing as well possibly in tandem with kiln temperature
Studies vary in terms of their intended breadth.

**Enumerative study** – intended to investigate a particular finite group of objects
- What is the compression strength of this batch of taconite?
- How many of the ball bearings in this shipment meet specifications?

**Analytical study** – investigate the general process. The intended scope is beyond the items at hand.
- How does drying temperature affect plywood strength? Not interested in just the particular boards on hand.
• **Population** – the universe of objects of interest
  All ball bearings in this shipment. Enumerative. Finite population.
  Strengths of plywood boards with this glue. Analytical. Unbounded, conceptually infinite population

• **Sample** – the items actually measured
  The ball bearing we take from the shipment and measure.
  The boards we glue up and then measure strengths
Example

Effects of freezing on glue bond strength

- Glue boards. Cut in half.
- Put half in freezer.
- Measure strength for each pair.
1.2.2 Types of Data

- **Qualitative/categorical** measurements or observations
  - Item identified by customer service call
  - Pass/ no pass compression strength test

- **Quantitative** measurements can be **counts** or **continuous**
  - Number of pellets failing compression test – count
  - Compression strengths of pellets – continuous (conceptually)
When possible, record the data at the most detailed level.

– From recorded compression strengths you can always determine number passing

– If you record only the number passing the test, you can't recover the compression strengths later
Data can be **univariate** or **multivariate**

- **Multivariate** – multiple measurements on the same unit
  - Brightness and strength of paper
    - 2 paired or bivariate data values

- **Univariate** – just one measurement
  - Just measure brightness
  - Measure brightness and strength on different sheets
  - Compression strengths for some pellets, iron content for others
Repeated measures data

• When multivariate data consist of several determinations of basically the same characteristic, the data are called repeated measures data.

• In the special case of bivariate responses, the term paired data is used.
If data are collected multivariately, retain this information in the data file.

– Don't create 2 unrelated files for strength and iron content losing info on which measurements were from the same pellet.

When possible take measurements multivariately

– Possible to investigate relationships
– Possible to reduce noise in observations
1.2.3 Types of Data Structures

A **complete factorial** study is one in which several process variables (and settings of each) are identified as being of interest, and data are collected under each possible combination of settings.

– The process variables are called the **factors**.
– The settings of a factor are its **levels**.
Paper airplanes

- Factor 1: Plane design
  Levels: design 1, design 2
- Factor 2: Paper weight
  Levels: light, medium, heavy

2x3 design (2 levels of factor 1, 3 levels of factor 2)

<table>
<thead>
<tr>
<th>Paper Weight</th>
<th>Light</th>
<th>Med</th>
<th>Heavy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design 1</td>
<td>____</td>
<td>____</td>
<td>____</td>
</tr>
<tr>
<td>Design 2</td>
<td>____</td>
<td>____</td>
<td>____</td>
</tr>
</tbody>
</table>

5 planes made for each treatment
30 planes in all (2x3x5)
Section 1.3 Measurement: Its Importance and Difficulty

- **Validity, measurement error, accuracy, precision**

See the book for these points and distinctions between them.

"It is impossible to overstate the importance of facing the question of measurement validity before plunging ahead in a statistical engineering study."
Variability is the root of the need for us to study statistical methods.

Variability is often from many sources including measurement error.

– Taconite pellets have different compression strengths.

– In addition even if we could measure compression strength on the same pellet twice, the answers won't come out the same.
Calibration, instrument drift

Example 8 - Different technicians used gauge differently.
   – Be sure to train technicians.
   – Run pilot studies.

Check on accuracy throughout. Instruments may drift over time.
   – calibration
   – Look at runs charts
"All models are false. Some models are useful." George Box

Read and heed the advice/information in the text.