1 Introduction

First we start with a road map for the course. The subject of **engineering statistics** is defined, its importance is described and basic terminology is introduced.

1.1 What is statistics?

Statistics is the science of collecting, presenting, analyzing, and making decisions from data. Often, as an engineer, it is necessary to **collect and interpret data** that will help in understanding how a new system or product works.

Statistics has applications to engineering through quality control, process control, reliability, risk management, system identification, design of experiments, etc.

Definition 1.1. *Engineering statistics* is the study of how best to

1. collect engineering data,
2. summarize or describe engineering data, and
3. draw formal inference and practical conclusions on the basis of engineering data,

all while recognizing the reality of variation.

We can break down this study into three main tasks:

1. **Summary:** describe, summarize and display data
2. **Inference:** draw conclusions from data
3. **Interpretation:** explain those conclusions in layman's terms (i.e. to people outside of statistics)

Example 1.1 (Heat treating gears, pg. 2). A process engineer is faced with the question, "How should gears be loaded into a continuous carburizing furnace in order to minimize distortion during heat treating?" The engineer conducts a well-thought-out study and obtains the runout values for 38 gears laid and 39 gears hung.

hung	laid
7, 8, 8, 10, 10,	5, 8, 8, 9, 9, 9,
10, 10, 11, 11,	9, 10, 10, 10,
11, 12, 13, 13,	11, 11, 11, 11,
13, 15, 17, 17,	11, 11, 11, 12,
17, 17, 18, 19,	12, 12, 12, 13,
19, 20, 21, 21,	13, 13, 13, 14,
21, 22, 22, 22,	14, 14, 15, 15,
23, 23, 23, 23,	15, 15, 16, 17,
24, 27, 27, 28,	17, 18, 19, 27
31, 36	

Table 1: Thrust face runouts (.0001 in.)

This data, as is, is hard to get insights from. Should the gears be hung or laid? We still cannot tell by looking at the table.

Numerical summaries:

mean laid runout: 12.6 (.0001 in)

mean hung runout: 17.9 (.0001 in)

Variation: even within a loading method

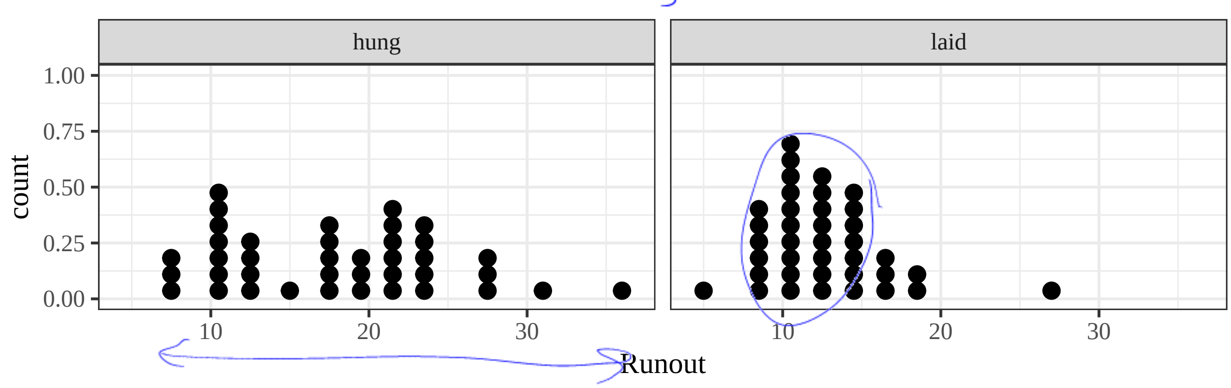


Figure 1: Dot diagrams of runouts.

hung gears more variable than laid gears

Inference: Runout values are variable, is there any assurance that the difference seen in the present data would reappear in further testing?

Points to need for methods of formal statistical inference from data and translation into practical conclusions. We can use statistical inference to say

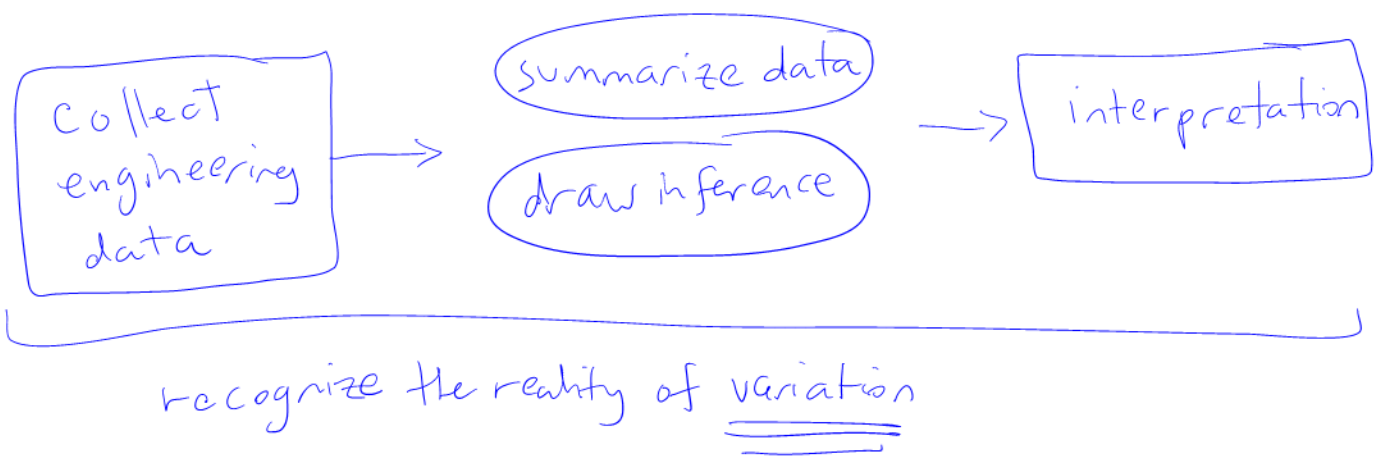
1. We are roughly 90% sure that the difference in long-run mean runout produced under conditions like those in the study is in the range 3.2 to 7.4.

2. We are roughly 95% sure that the difference in long-run mean runout produced under conditions like those in the study is in the range 3.0 to 22.2.

Interpretation: Explains those answers in layman's terms without all the probability theory, symbols and equations.

Combine with other information, such as consequences of a given amount of runout and the cost for hanging and laying gears $\implies$ apply sound engineering judgement.

Overview/schematic:



1.2 Basic terminology

It's important we speak the same language. This section introduces common terminology related to statistical studies, types of data, and types of data structures.

1.2.1 Population vs. sample

Definition 1.2. A *population* is the entire group of objects about which one wishes to gather information in a statistical study.

Definition 1.3. A *sample* is the group of objects on which one actually gathers data.

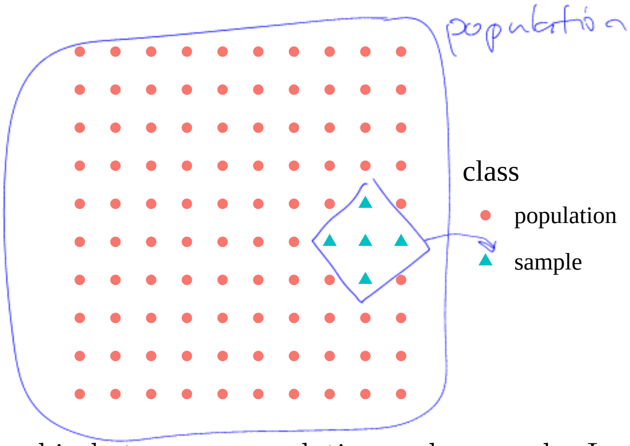


Figure 2: The relationship between a population and a sample. In this example, we have 100 parts and 5 are examined in order to verify acceptibility. Notice we say "one sample", not "five samples".

Example 1.2 (Heat treating gears, cont'd)**.**

Population: All the gears with the same make and model as those included in the experiment.

Sample: The 77 gears arranged, tested, and measured for distortion.

Example 1.3 (Fiscal cliff)**.** On Dec. 1-2, 2012, Gallup conducted a study to find out what proportion of Americans prefer a compromise on the Fiscal Cliff issue. 1000 adults were randomly selected for telephone interviews. The adults were aged 18 and older and living in any of the 50 U.S. states or the District of Columbia.

Population: All Americans in the 50 states + DC aged 18 and older.

Sample: 1000 adults who were called.

Example 1.4. Esbit manufactures fuel pellets out of compressed hexamine powder. Suppose a new shipment of 100 pelletizing machines arrives, and the goal of a new study is to determine the quality of this particular new shipment.

5 machines out of the 100 are randomly selected for comprehensive testing in which each produces 200 pellets, and each pellet's mass, volume, flash point, and rate of combustion are measured.

Population: 100 new pelletizing machines in the shipment.

Sample: 5 machines selected for testing.

1.2.2 Types of studies

When an engineer **collects data**, (s)he must decide how active to be. Should the engineer manipulate the process or let things happen and record the results?

Definition 1.4. An *experimental study* (or, more simply, an experiment) is one in which the investigator's role is active. Process variables are manipulated, and the study environment is regulated.

Definition 1.5. In a *randomized experiment*, investigators control the assignment of treatments to experimental units using a chance mechanism (like the flip of a coin or a computer's random number generator). "random" = "chance"

Example 1.5 (Smoking - experimental)**.** To answer the question of if smoking affects lung capacity in young men, an investigator does the following.

Finds 100 men age 20 who do not currently smoke. Randomly assigns 50 of the 100 men to the smoking treatment and the other 50 to the non-smoking treatment. Those in the smoking group smoke a pack a day for 10 years while those in the control group remain smoke free for 10 years. Measures lung capacity for each of the 100 men. Analyze, interpret, and draw conclusions from data.

Definition 1.6. An *observational study* is one in which the investigator's role is basically passive. A process or phenomenon is watched and data are recorded, but there is no intervention on the part of the person conducting the study.

Example 1.6 (Smoking - observational)**.** To answer the question of if smoking affects lung capacity in young men, an investigator does the following.

Finds 100 men age 30 of which 50 have been smoking a pack a day for 10 years while the other 50 have been smoke free for 10 years. Measures lung capacity for each of the 100 men. Analyze, interpret, and draw conclusions from data.

1.2.3 Data

Engineers encounter many types of data. It's important to have a way to distinguish all the different types of data you will see.

Definition 1.7. *Qualitative* or *categorical* data are the values of basically nonnumerical characteristics associated with items in a sample. There can be an order to qualitative data, but aggregation and counting are required to produce meaningful numeric values from such data.

Definition 1.8. *Quantitative* or *numeric* data are the values of numerical characteristics associated with items in a sample. These are typically counts of the number of occurrences of a phenomenon of interest or measurements of some physical property of the items.

Types of numeric data:

- **Discrete:** measurements are separated points (e.g. # of pages in a book)
- **Continuous:** measurements lie in a continuum (e.g. mpg of a car)

Definition 1.9. *Univariate data* arise when only a single characteristic of each sampled item is observed.

Definition 1.10. *Multivariate data* arise when observations are made on more than one characteristic of each sampled item. A special case is when there are two characteristics - *bivariate*.

Example 1.7 (Heat treating gears, cont'd)**.**

hung	laid
7, 8, 8, 10, 10,	5, 8, 8, 9, 9, 9,
10, 10, 11, 11,	9, 10, 10, 10,
11, 12, 13, 13,	11, 11, 11, 11,
13, 15, 17, 17,	11, 11, 11, 12,
17, 17, 18, 19,	12, 12, 12, 13,
19, 20, 21, 21,	13, 13, 13, 14,
21, 22, 22, 22,	14, 14, 15, 15,
23, 23, 23, 23,	15, 15, 16, 17,
24, 27, 27, 28,	17, 18, 19, 27
31, 36	

Table 2: Thrust face runouts (.0001 in.)

Univariate or **bivariate**?

Arrange the data in a table, where:

- Each row is a sample unit, or thing that you measure (gear, in this case).
- Each column is a variable, or characteristic that you control or measure.

Arrangement	Runout
hung	21
hung	8
hung	20
laid	11
laid	16
laid	10
...	...

Table 3: Thrust face runouts (.0001 in.) rearranged.

Definition 1.11. When multivariate data consist of several determinations of basically the same characteristic (e.g., made with different instruments or at different times), the data are called *repeated measures data*. In the special case of bivariate responses, the term *paired data* is used.

Example 1.8 (Paired distortion)**.** For the gears heat treating example, the measurements were actually made on the 77 gears both before and after heat treating.

1.2.4 Data structures

It is common for several sets of conditions to be compared with each other, in which **several samples** are involved. Here are two structures for multisample studies.

Definition 1.12. A *response variable* (or dependent variable) is the outcome of a study.

Definition 1.13. A *factor* is any numerical or categorical variable with a finite set of possible values. A *level* is the value of the factor.

Definition 1.14. 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 of the process variables. The process variables are usually factors.

Example 1.9 (Pelletizing machine, pg. 6, 12)**.** Experimentation with a pelletizing machine using a $2 \times 2 \times 2$ or 2^3 factorial structure. The researchers are measuring the percentage of acceptable fuel pellets for various situations. The factors and respective levels are:

- *Die volume* - low volume vs. high volume (1 factor w/ 2 levels)
- *Material flow* - current method vs. manual filling (1 factor w/ 2 levels)
- *Mixture type* - no binding agent vs. with binder (1 factor w/ 2 levels)

$2 \times 2 \times 2 = 2^3 = 8$

There are then 8 sets of conditions under which data are collected.

Volume	Flow	Mixture
low	current	no bunder
high	current	no bunder
low	manual	no bunder
high	manual	no bunder
low	current	binder
high	current	binder
low	manual	binder
high	manual	binder

Table 4: Combinations in a 2^3 factorial study.

When there are many factors or levels are involved, the number of sampling units in a complete fractional study can quickly reach an impractical size.

Definition 1.15. A *fractional factorial study* is one in which data are collected for only some of the combinations that would make up a complete factorial study.

Example 1.10 (Pelletizing machine, cont'd)**.**

Volume	Flow	Mixture
high	current	no bunder
low	manual	no bunder
low	current	binder
high	manual	binder

Table 5: A "clever" half of the eight possible combinations.

→ 8.3, 8.4 of text

1.3 Measurement difficulties

A measurement system is:

- **Valid** if usefully or appropriately represents the feature of an object/system of engineering importance
- **Accurate** if on average it produces the true or correct value of a quantity being measured ("unbiased")
- **Precise** if it produces small variation in repeated measures of the same object

Example 1.11 (Paper thickness, pg. 16)**.** Two students measured the thickness of paper in the same book. Each student took 10 measurements.

Gulliver	Wendel
0.0972, 0.0964,	0.0807, 0.0826,
0.0978, 0.0971,	0.0854, 0.0817,
0.096, 0.0947,	0.0824, 0.0799,
0.12, 0.0991,	0.0812, 0.0807,
0.098, 0.1033	0.0816, 0.0804

Table 6: Thickness of book paper in mm.

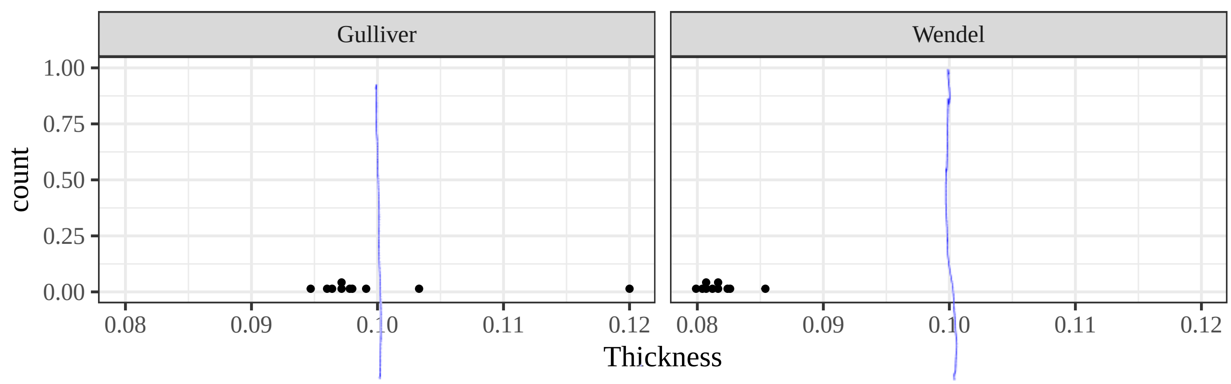


Figure 3: Dot diagrams of paper thickness.

If the true thickness is 0.1, then

1. Who is more accurate? Gulliver (measurements center ~ 0.10mm)
2. Who is more precise? Wendel (measurements are consistent, less varied)