# Engineering Probability & Statistics
# (AGE 1150)
# Chapter 1: Introduction

Dr. O. Philips Agboola

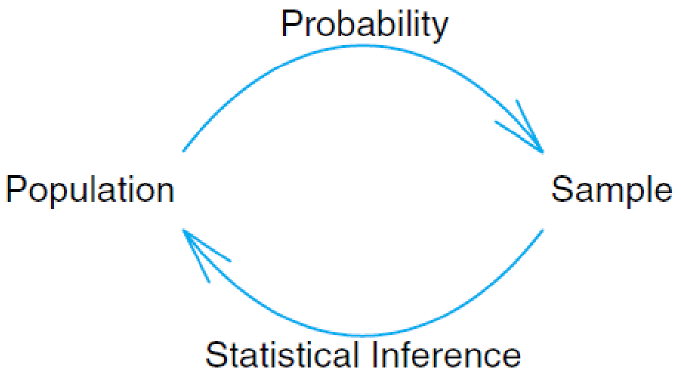
# Chapter 1: Introduction

- 1- the Japanese “industrial miracle,” which began in the middle of the 20th century. The Japanese were able to succeed where other countries had failed–namely, to create an atmosphere that allows the production of high-quality products. Why?
- Ans: Much of the success of the Japanese has been attributed to the use of statistical methods and statistical thinking among management personnel.
- scientific data vs inferential statistics. These statistical methods are designed to contribute to the process of making scientific judgments in the face of uncertainty and variation.
- Statistical methods are used to analyze data from a process in order to gain more sense of where in the process changes may be made to improve the quality of the process.

- Statistical methods are used to analyze data from a process such as this one in order to gain more sense of where in the process changes may be made to improve the qualityof the process.
- System population (population is collection of all individuals or individual items of a particular type).
- Scientific systems information is gathered in the form of samples, or collections of observations.
- These samples are taken from population.

# Probability And Statistics

- Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1 (where 0 indicates impossibility and 1 indicates certainty).
- Event is a set of outcomes of an experiment.
- Elements of probability allow us to quantify the strength or confidence in our conclusions.
- Statistics is the study of the collection, analysis, interpretation, presentation, and organization of data.




Fundamental relationship between probability and inferential statistics.

Table 1.1: Data Set for Example 1.2

| **No Nitrogen** | **Nitrogen** |
|---|---|
| 0.32 | 0.26 |
| 0.53 | 0.43 |
| 0.28 | 0.47 |
| 0.37 | 0.49 |
| 0.47 | 0.52 |
| 0.43 | 0.75 |
| 0.36 | 0.79 |
| 0.42 | 0.86 |
| 0.38 | 0.62 |
| 0.43 | 0.46 |

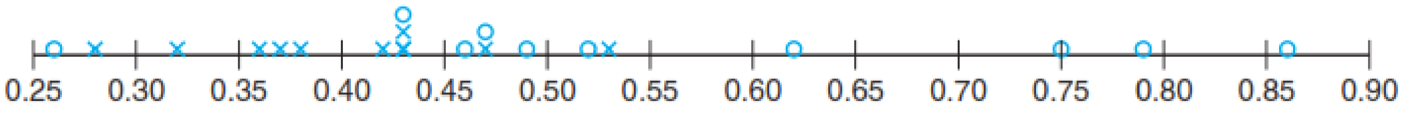



Figure 1.1: A dot plot of stem weight data.

# Measures of Location

- Sample size (n), means the number of elements in the sample
- Mean (Average)

Suppose that the observations in a sample are $x_1, x_2, \ldots, x_n$. The **sample mean**, denoted by $\bar{x}$, is

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}.$$

- Median

Given that the observations in a sample are $x_1, x_2, \ldots, x_n$, arranged in **increasing order** of magnitude, the sample median is

$$\tilde{x} = \begin{cases} x_{(n+1)/2}, & \text{if } n \text{ is odd,} \\ \frac{1}{2}(x_{n/2} + x_{n/2+1}), & \text{if } n \text{ is even.} \end{cases}$$

Example:
Calculate the sample mean and median for the following data set:
1.7, 2.2, 3.11, 3.9, and 14.7.

- n = 5 and n is odd
- Mean is

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

= (1.7 + 2.2 + 3.9 + 3.11 + 14.7 ) / 5 = 5.12

- Median is

$$\tilde{x} = \begin{cases} x_{(n+1)/2}, & \text{if } n \text{ is odd,} \\ \frac{1}{2}(x_{n/2} + x_{n/2+1}), & \text{if } n \text{ is even.} \end{cases}$$

- n is odd then the median is the x value have the arrange (n+1)/2 = (5+1)/2 = 3 after arranging the values from the smallest to the largest. i.e. the third element which is 3.11.

# Measure of Variability

- Sample range: The simplest measure of variability or spread. It is $X_{max} - X_{min}$
- Standard deviation: measure of spread or variability often used mostly. $x_1, x_2, \ldots, x_n$ denote sample values, then:

The **sample variance**, denoted by $s^2$, is given by

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}.$$

The **sample standard deviation**, denoted by $s$, is the positive square root of $s^2$, that is,

$$s = \sqrt{s^2}.$$

- The quantity $(n - 1)$ is called the degrees of freedom associated the variance.

Example:
Find (compute) the sample variance and standard deviation of the data set (5, 17, 6, 4).

- First compute the mean value (or average) as before
- Average = 8 (check this yourself)
- Degree of freedom = n -1 = 4 – 1 = 3
- $s^2 =$ $(5-8)^2 + (17-8)^2 + (6-8)^2 + (4-8)^2 = (-3)^2 + 9^2 + (-2)^2 + (-4)^2$/3
- variance ($s^2$)= 110 / 3 = 36.67
- Standard deviation (s) = 6.06

Example:
an engineer is interested in testing the "bias" in a pH meter. Data are collected on the meter by measuring the pH of a neutral substance (pH = 7.0). A sample of size 10 is taken, with results given by
7.07 7.00 7.10 6.97 7.00 7.03 7.01 7.01 6.98 7.08.

- The sample mean is given by

$$\bar{x} = \frac{7.07 + 7.00 + 7.10 + \cdots + 7.08}{10} = 7.0250.$$

- The sample variance ($s^2$) is given by

$$s^2 = \frac{1}{9}[(7.07 - 7.025)^2 + (7.00 - 7.025)^2 + (7.10 - 7.025)^2 + \cdots + (7.08 - 7.025)^2] = 0.001939.$$

- The sample standard deviation (s) is given by

$$s = \sqrt{0.001939} = 0.044$$

# Homework 1

**1.1** The following measurements were recorded for the drying time, in hours, of a certain brand of latex paint.

| | | | | |
|---|---|---|---|---|
| 3.4 | 2.5 | 4.8 | 2.9 | 3.6 |
| 2.8 | 3.3 | 5.6 | 3.7 | 2.8 |
| 4.4 | 4.0 | 5.2 | 3.0 | 4.8 |

Assume that the measurements are a simple random sample.

(a) What is the sample size for the above sample?

(b) Calculate the sample mean for these data.

(c) Calculate the sample median.

(d) Plot the data by way of a dot plot.

**1.3** A certain polymer is used for evacuation systems for aircraft. It is important that the polymer be resistant to the aging process. Twenty specimens of the polymer were used in an experiment. Ten were assigned randomly to be exposed to an accelerated batch aging process that involved exposure to high temperatures for 10 days. Measurements of tensile strength of the specimens were made, and the following data were recorded on tensile strength in psi:

| | | | | | |
|---|---|---|---|---|---|
| No aging: | 227 | 222 | 218 | 217 | 225 |
| | 218 | 216 | 229 | 228 | 221 |
| Aging: | 219 | 214 | 215 | 211 | 209 |
| | 218 | 203 | 204 | 201 | 205 |

(a) Do a dot plot of the data.

(b) From your plot, does it appear as if the aging process has had an effect on the tensile strength of this

**1.7** Consider the drying time data for Exercise 1.1 on page 13. Compute the sample variance and sample standard deviation.

# Homework 1

**1.9** Exercise 1.3 on page 13 showed tensile strength data for two samples, one in which specimens were exposed to an aging process and one in which there was no aging of the specimens.

(a) Calculate the sample variance as well as standard deviation in tensile strength for both samples.

(b) Does there appear to be any evidence that aging affects the variability in tensile strength? (See also the plot for Exercise 1.3 on page 13.)