Π

Lectures' Notes

STAT – 324
Probability and Statistics for Engineers

Summer Semester 1426/1427

Instructor: Dr. Abdullah Al-Shiha

Office: 2B43 Building # 4
Office Hours: To be announced

Department of Statistics and Operations Research
College of Science, King Saud University

Textbook:
Probability and Statistics for Engineers and Scientists
By: R. E. Walpole, R. H. Myers, and S. L. Myers
(Sixth Edition or beyond)

Chapter 1: Introduction to Statistics and Data Analysis:

1.1 Introduction:

* Populations and Samples:

Population	→	Sample = Observations
(Some Unknown Parameters)		**(We calculate Some Statistics)**
Example: KSU Students		Example: 20 Students from KSU
(Height Mean)		**(Sample Mean)**
N=Population Size		**n = Sample Size**

- Let $X_1, X_2, \ldots, X_N$ be the population values (in general, they are unknown)
- Let $x_1, x_2, \ldots, x_n$ be the sample values (these values are known)
- Statistics obtained from the sample are used to estimate (approximate) the parameters of the population.

* Scientific Data

* Statistical Inference

(1) Estimation:

$\rightarrow$ Point Estimation

$\rightarrow$ Interval Estimation (Confidence Interval)

(2) Hypotheses Testing

1.3 Measures of Location (Central Tendency):

- The data (observations) often tend to be concentrated around the center of the data.
- Some measures of location are: the mean, mode, and median.
- These measures are considered as representatives (or typical values) of the data. They are designed to give some quantitative measures of where the center of the data is in the sample.

The Sample mean of the observations ($\bar{x}$):

If $x_1, x_2, \ldots, x_n$ are the sample values, then the sample mean is

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \text{(unit)}$$

Example:

Suppose that the following sample represents the ages (in year) of a sample of 3 men:

$$x_1 = 30,\ x_2 = 35,\ x_3 = 27. \quad (n=3)$$

Then, the sample mean is:

$$\bar{x} = \frac{30+35+27}{3} = \frac{92}{3} = 30.67 \quad \text{(year)}$$

Note: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$.

1.4 Measures of Variability (Dispersion or Variation):

- The variation or dispersion in a set of data refers to how spread out the observations are from each other.
- The variation is small when the observations are close together. There is no variation if the observations are the same.
- Some measures of dispersion are range, variance, and standard deviation
- These measures are designed to give some quantitative measures of the variability in the data.

The Sample Variance (S^2):

Let $x_1, x_2, \ldots, x_n$ be the observations of the sample. The sample variance is denoted by S^2 and is defined by:

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n-1} \quad (\text{unit})^2$$

where $\bar{x} = \sum_{i=1}^{n} x_i / n$ is the sample mean.

Note:

(n −1) is called the degrees of freedom (df) associated with the sample variance S^2.

The Standard Deviation (S):

The standard deviation is another measure of variation. It is the square root of the variance, i.e., it is:

$$S=\sqrt{S^2}=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \quad \text{(unit)}$$

Example:

Compute the sample variance and standard deviation of the following observations (ages in year): 10, 21, 33, 53, 54.

Solution:

n=5

$$\bar{x}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{\sum_{i=1}^{5}x_i}{5}=\frac{10+21+33+53+54}{5}=\frac{171}{5}=34.2 \quad \text{(year)}$$

$$S^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}=\frac{\sum_{i=1}^{5}(x_i-34.2)^2}{5-1}$$

$$=\frac{(10-34.2)^2+(21-34.2)^2+(33-34.2)^2+(53-34.2)^2+(54-34.2)^2}{4}$$

$$=\frac{1506.8}{4}=376.7 \quad \text{(year)}^2$$

The sample standard deviation is:

$$S=\sqrt{S^2}=\sqrt{376.7}=19.41 \qquad \text{(year)}$$

* Another Formula for Calculating S^2:

$$S^2=\frac{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}{n-1} \qquad \text{(It is simple and more accurate)}$$

For the previous Example,

x_i	10	21	33	53	54	$\sum x_i=171$
x_i^2	100	441	1089	2809	2916	$\sum x_i^2=7355$

$$S^2=\frac{\sum_{i=1}^{n}x_i^2-n\bar{x}^2}{n-1}=\frac{7355-(5)(34.2)^2}{5-1}=\frac{1506.8}{4}=376.7 \quad \text{(year)}^2$$