

## **Note of Math Statistics**

## **Course Note**

2024-11-04

Lawrence

2024 Year Fall Season Class

## 1 Introduction

## 1.1 Definition

We set a population F and  $X_1, X_2, ..., X_n$  is a random sample of size n from the population F. If  $X_1, X_2, ..., X_n$  are independent and identically distributed random variables, (i.d.d)then their common probability distributed function is:

$$F(x_1)F(x_2)...F(x_n) = \prod_{i=1}^n F(x_i)$$
 (1.1.1)

Their common probability density function is:

$$f(x_1)f(x_2)...f(x_n) = \prod_{i=1}^n f(x_i)$$
 (1.1.2)

**Statistic** Given a random sample  $X_1, X_2, ..., X_n$  from a population F, a statistic is any real function  $g(X_1, X_2, ..., X_n)$  of the sample values  $X_1, X_2, ..., X_n$ .

Sample mean The sample mean is the statistic defined by:

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \tag{1.1.3}$$

Sample variance The sample variance is the statistic defined by:

$$S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} \left( X_{i} - \overline{X} \right)^{2}$$
 (1.1.4)

Sample standard deviation The sample standard deviation is the statistic defined by:

$$S = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2} j$$
 (1.1.5)

They have the following properties:

$$\sum_{i=1}^{n} \left( X_i - \overline{X} \right) = 0 \tag{1.1.6}$$

2. We apply a transforation to sample  $Y_i = aX_i + b$ . Then the sample mean and sample variance of  $Y_i$  are:

$$\overline{Y} = a\overline{X} + b \tag{1.1.7}$$

$$S_Y^2 = a^2 S_X^2 (1.1.8)$$

1. To any constant c, we have:

$$\sum_{i=1}^{n} (X_i - c)^2 \ge \sum_{i=1}^{n} (X_i - \overline{X})^2$$
 (1.1.9)

**Sample moment** The k-th sample origin moment is the statistic defined by:

$$a_{n,k} = \frac{1}{n} \sum_{i=1}^{n} X_i^k \tag{1.1.10}$$

The k-th sample central moment is the statistic defined by:

$$m_{n,k} = \frac{1}{n} \sum_{i=1}^{n} \left( X_i - \overline{X} \right)^k \tag{1.1.11}$$

**Order statistics** Given a random sample  $X_1, X_2, ..., X_n$  from a population F, we arrange the sample values in increasing order: $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(n)}$ . We call  $X_{(1)}, X_{(2)}, ..., X_{(n)}$  the order statistics of the sample.

Sample median The sample median is the statistic defined by: