# Chapter 6 - Lecture 2
# The distribution of a linear combination

Yuan Huang

January 11th, 2013

## Definition of Linear Combination

We have

- a random sample $X_1, X_2, \ldots, X_n$
- $n$ constants $a_1, \ldots, a_n$

then the random variable

$$Y = a_1X_1 + \ldots + a_nX_n = \sum_{i=1}^{n} a_iX_i \tag{1}$$

is called a **linear combination of $X$'s**.

Many statistics are linear functions of the sample data $X_1, \ldots, X_n$:

$$Y = a_1 X_1 + \ldots + a_n X_n = \sum_{i=1}^{n} a_i X_i.$$

1. $\bar{X} = \frac{1}{n} X_1 + \ldots + \frac{1}{n} X_n$;

By learning properties of linear combination, we can get a clearer view of how a statistic is distributed.

# For general sample

**Proposition 1**

$$E(\sum_{i=1}^{n} a_i X_i) = a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n).$$

- This proposition holds no matter whether the $X_i$'s are independent or not.
- Interpretation, the sampling distribution of $\sum_{i=1}^{n} a_i X_i$ has mean $a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n)$.
- In most general case, each $X_i$ has expectation $\mu_i$, then

$$E(\sum_{i=1}^{n} a_i X_i) = a_1 \mu_1 + a_2 \mu_2 + \ldots + a_n \mu_n$$

Practice: If $E(X_1) = 2$ and $E(X_2) = 3$ and $E(X_3) = 1$ then

- $E(X_1 - X_2)$?
- $E(X_1 + X_2 - X_3)$?
- $E(\bar{X})$?

# For general sample

**Proposition 2:**

$$V(a_1X_1 + \ldots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j).$$

# For general sample

**Proposition 2:**

$$V(a_1X_1 + \ldots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j).$$

Note that

1. $Cov(X_i, X_i) = V(X_i)$;
2. If $X_i$ and $X_j$ are independent, $Cov(X_i, X_j) = 0$ (uncorrelated);

# For general sample

**Proposition 2:**

$$V(a_1X_1 + \ldots + a_nX_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j).$$

Note that

1. $Cov(X_i, X_i) = V(X_i)$;
2. If $X_i$ and $X_j$ are independent, $Cov(X_i, X_j) = 0$ (uncorrelated);

## Corollary

*If $X_1, \ldots, X_n$ are mutually independent, then*

$$V(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 V(X_i).$$

# Special case for iid random sample

We have a random sample $X_1, X_2, \ldots, X_n$ from a distribution with mean $\mu$ and variance $\sigma^2$, and let $Y = \sum_{i=1}^n a_i X_i$ then:

$$\mu_Y = \sum_{i=1}^n a_i \mu = \mu \sum_{i=1}^n a_i$$

and

$$\sigma_Y^2 = \sum_{i=1}^n a_i^2 \sigma^2 = \sigma^2 \sum_{i=1}^n a_i^2$$

# Example

If we have $X_1$ and $X_2$ that $X_1$ has mean $\mu_1$ and variance $\sigma_1^2$, $X_2$ has mean $\mu_2$ and variance $\sigma_2^2$

1. What is $E(X_1 + X_2)$ and $V(X_1 + X_2)$, when
   - If $X_1, X_2$ are independent:
   - If $X_1, X_2$ are dependent:
2. What is $E(X_1 - X_2)$ and $V(X_1 - X_2)$, when
   - If $X_1, X_2$ are independent:
   - If $X_1, X_2$ are dependent:

## Examples

# Example 6.11 page 301

A gas station sells three grades of gasoline: regular unleaded, extra unleaded, and super unleaded. These are priced at $2.20, $2.35, $2.50 per gallon, respectively. Let $X_1$, $X_2$ and $X_3$ denote the amounts of these grades purchased (gallons) on a particular day. Suppose the $X_i's$ are independent with $\mu_1 = 1000$, $\mu_2 = 500$, $\mu_3 = 300$, $\sigma_1 = 100$, $\sigma_2 = 80$ and $\sigma_3 = 50$. The revenue from sales is $Y = 2.2X_1 + 2.35X_2 + 2.5X_3$.

1. What is $E(Y)$?
2. What is $V(Y)$?

# For general normal random variables

**Proposition 3**

When $X_1, \ldots, X_n$ are independent and normally distributed, suppose $X_i \sim N(\mu_i, \sigma_i^2)$, then for any linear combination $Y = a_1X_1 + \ldots + a_nX_n = \sum_{i=1}^n a_iX_i$,

$$Y \sim N(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2).$$

Remark. This proposition is true ONLY for Normal Random Variables.

# For iid normal random sample

## Corollary

*When $X_1, \ldots, X_n$ are i.i.d and $X_i \sim N(\mu, \sigma^2)$, then for any linear combination $Y = a_1X_1 + \ldots + a_nX_n = \sum_{i=1}^{n} a_iX_i$,*

$$Y \sim N((\sum_{i=1}^{n} a_i)\mu, (\sum_{i=1}^{n} a_i^2)\sigma^2).$$

# Introduce another tool to derive distribution

**Proposition 4**

Let $X_1, X_2, \ldots, X_n$ independent random variables with mgfs $M_{X_i}(t)$ and $Y$ is the linear combination defined in equation (1), then

$$M_Y(t) = M_{X_1}(a_1 t) \times M_{X_2}(a_2 t) \times \ldots \times M_{X_n}(a_n t) \qquad (2)$$

# Normal case

$X$ and $Y$ are independent Normal random variable. $X$ has mean $\mu_1$ and variance $\sigma_1$. $Y$ has mean $\mu_2$ and variance $\sigma_2$. What's the distribution of $X + Y$ ?

# Poisson case

$X$ and $Y$ are independent Poisson random variable. $X$ has mean $\nu$ and $Y$ has mean $\lambda$. What's the distribution of $X + Y$ ? ( Example 6.16 page 306)

Homework for Section 6.3: 33, 34, 44.

## HW1

- Due next Jan. 18
- Hand-in: (Sec 6.1 P290) 2, 3 ; (Sec 6.3 P306) 33, 34, 44
- Not-Hand-in: Reading
  1. Book sections 6.1, 6.3
  2. [Reading 1] under Readings tag of course website.