Probability and Statistics COSE112

College of Informatics
Korea University

Chapter 4

Mathematical Expectation

Section 4.1

Mean of a
Random Variable

Definition 4.1

Let X be a random variable with probability distribution $f(x)$. The **mean**, or **expected value**, of X is

$$\mu = E(X) = \sum_x x f(x)$$

if X is discrete, and

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) \, dx$$

if X is continuous.

- What is the expected number of H for two tosses of a fair coin?

Example 4.1

Example 4.1: A lot containing 7 components is sampled by a quality inspector; the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Example 4.3

Example 4.3: Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \frac{20,000}{x^3}, & x > 100, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected life of this type of device.

Theorem 4.1

Let X be a random variable with probability distribution $f(x)$. The expected value of the random variable $g(X)$ is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x)$$

if X is discrete, and

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)\ dx$$

if X is continuous.

Example 4.5

Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of $g(X) = 4X + 3$.

Probability is Expected Indictor

- indicator function is given by

$$1_A(x) = \begin{cases} 1 & x \in A \\ 0 & \text{otherwise} \end{cases}$$

- We have

$$\mathbb{E}[1_A(X)] = \mathbb{P}(X \in A)$$

- for example,

$$\mathbb{P}(a \leq X \leq b) = \mathbb{E}[1_{(a,b)}(X)] = \int_a^b f(x)dx$$

Definition 4.2

Let X and Y be random variables with joint probability distribution $f(x, y)$. The mean, or expected value, of the random variable $g(X, Y)$ is

$$\mu_{g(X,Y)} = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

if X and Y are discrete, and

$$\mu_{g(X,Y)} = E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) \; dx \; dy$$

if X and Y are continuous.

Example 4.6

Example 4.6

Example 4.6: Let X and Y be the random variables with joint probability distribution indicated in Table 3.1 on page 116. Find the expected value of $g(X, Y) = XY$. The table is reprinted here for convenience.

	$f(x,y)$	x			Row Totals
		0	1	2	
	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
y	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Example 4.7

Example 4.7: Find $E(Y/X)$ for the density function

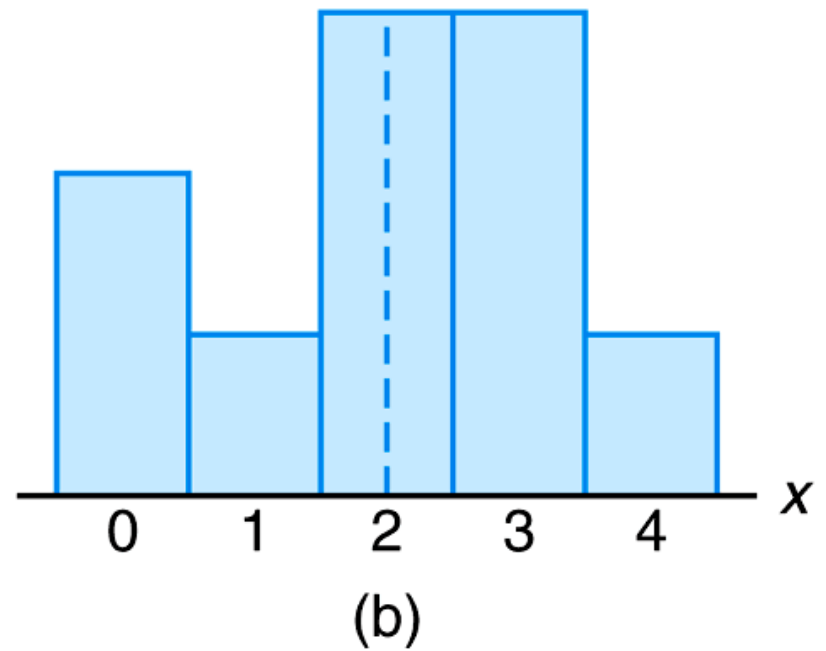$$f(x,y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2,\ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$
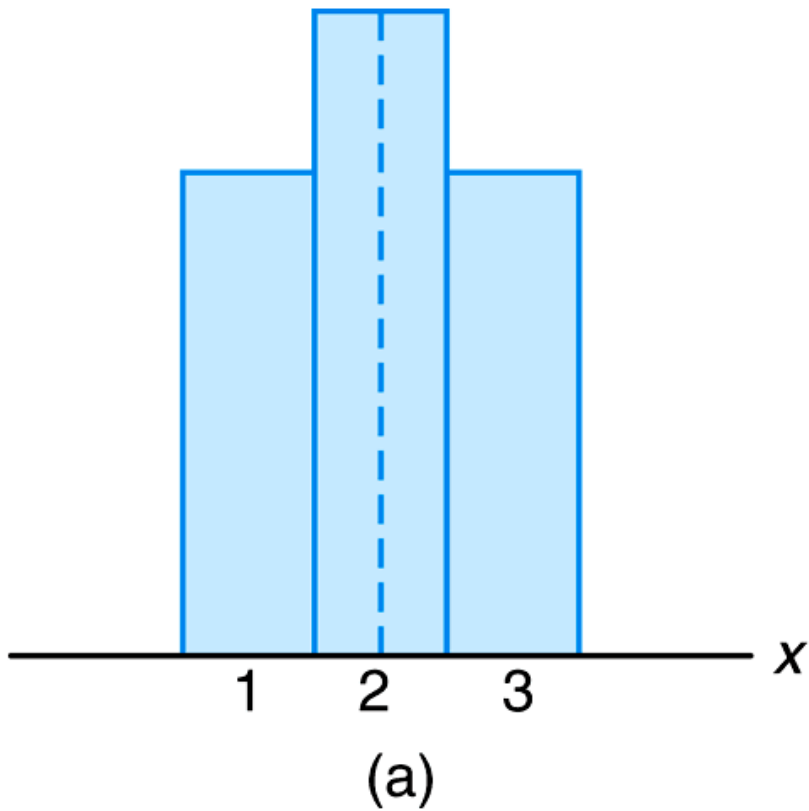
Solution: We have

$$E\left(\frac{Y}{X}\right) = \int_0^1 \int_0^2 \frac{y(1+3y^2)}{4}\, dxdy = \int_0^1 \frac{y+3y^3}{2}\, dy = \frac{5}{8}.$$

Section 4.2

Variance and Covariance of Random Variables

Figure 4.1 Distributions with equal means and unequal dispersions

(a)

(b)

Definition 4.3

Let X be a random variable with probability distribution $f(x)$ and mean μ. The variance of X is

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x), \qquad \text{if } X \text{ is discrete, and}$$

$$\sigma^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) \, dx, \qquad \text{if } X \text{ is continuous.}$$

The positive square root of the variance, σ, is called the **standard deviation** of X.

Theorem 4.2

The variance of a random variable X is

$$\sigma^2 = E(X^2) - \mu^2.$$

Example 4.9

Example 4.9: Let the random variable X represent the number of defective parts for a machine when 3 parts are sampled from a production line and tested. The following is the probability distribution of X.

x	0	1	2	3
$f(x)$	0.51	0.38	0.10	0.01

Using Theorem 4.2, calculate σ^2.

Example 4.10

Example 4.10: The weekly demand for a drinking-water product, in thousands of liters, from a local chain of efficiency stores is a continuous random variable X having the probability density

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the mean and variance of X.

Theorem 4.3

Let X be a random variable with probability distribution $f(x)$. The variance of the random variable $g(X)$ is

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \sum_x [g(x) - \mu_{g(X)}]^2 f(x)$$

if X is discrete, and

$$\sigma^2_{g(X)} = E\{[g(X) - \mu_{g(X)}]^2\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(X)}]^2 f(x)\ dx$$

if X is continuous.

Definition 4.4

Let X and Y be random variables with joint probability distribution $f(x, y)$. The covariance of X and Y is

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \sum_x \sum_y (x - \mu_X)(y - \mu_y)f(x, y)$$

if X and Y are discrete, and

$$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_y)f(x, y) \ dx \ dy$$

if X and Y are continuous.

Definition 4.4

- Covariance
  - Measure of how two RVs are 'aligned' when they change
  - Cov(X,Y)
    - positive: tend to vary in same direction
    - negative: tend to vary in opposite direction
    - If independent, Cov is zero
    - Cov is zero => does not mean independence

Theorem 4.4

The covariance of two random variables X and Y with means μ_X and μ_Y, respectively, is given by

$$\sigma_{XY} = E(XY) - \mu_X \mu_Y.$$

Example 4.14

Example 4.14: The fraction X of male runners and the fraction Y of female runners who compete in marathon races are described by the joint density function

$$f(x, y) = \begin{cases} 8xy, & 0 \leq y \leq x \leq 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the covariance of X and Y.

Definition 4.5

Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Definition 4.5

Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

- Can be shown that
$$-1 \le \rho_{XY} \le 1$$
- 1 or -1 if Y=aX+b (linear relation)

Definition 4.5

Let X and Y be random variables with covariance σ_{XY} and standard deviations σ_X and σ_Y, respectively. The correlation coefficient of X and Y is

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}.$$

Example 4.15: Find the correlation coefficient between X and Y in Example 4.13.

Section 4.3

Means and Variances of Linear Combinations of Random Variables

Theorem 4.5

If a and b are constants, then

$$E(aX + b) = aE(X) + b.$$

Corollary 4.1

Setting $a = 0$, we see that $E(b) = b$.

Corollary 4.2

Setting $b = 0$, we see that $E(aX) = aE(X)$.

Example 4.19

The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)].$$

Example 4.20

Example 4.20

Example 4.20: The weekly demand for a certain drink, in thousands of liters, at a chain of convenience stores is a continuous random variable $g(X) = X^2 + X - 2$, where X has the density function

$$f(x) = \begin{cases} 2(x - 1), & 1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Find the expected value of the weekly demand for the drink.

Theorem 4.7

The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions. That is,

$$E[g(X,Y) \pm h(X,Y)] = E[g(X,Y)] \pm E[h(X,Y)].$$

Corollary 4.3

Setting $g(X, Y) = g(X)$ and $h(X, Y) = h(Y)$, we see that

$$E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)].$$

Corollary 4.4

Setting $g(X, Y) = X$ and $h(X, Y) = Y$, we see that

$$E[X \pm Y] = E[X] \pm E[Y].$$

Theorem 4.8

Let X and Y be two independent random variables. Then

$$E(XY) = E(X)E(Y).$$

Corollary 4.5

Let X and Y be two independent random variables. Then $\sigma_{XY} = 0$.

Example 4.21

Example 4.21: It is known that the ratio of gallium to arsenide does not affect the functioning of gallium-arsenide wafers, which are the main components of microchips. Let X denote the ratio of gallium to arsenide and Y denote the functional wafers retrieved during a 1-hour period. X and Y are independent random variables with the joint density function

$$f(x, y) = \begin{cases} \frac{x(1+3y^2)}{4}, & 0 < x < 2, \ 0 < y < 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Show that $E(XY) = E(X)E(Y)$,

Theorem 4.9

If X and Y are random variables with joint probability distribution $f(x, y)$ and a, b, and c are constants, then

$$\sigma^2_{aX+bY+c} = a^2 \sigma^2_X + b^2 \sigma^2_Y + 2ab\sigma_{XY}.$$

Corollary 4.6

Setting $b = 0$, we see that

$$\sigma^2_{aX+c} = a^2 \sigma^2_X = a^2 \sigma^2.$$

Corollary 4.7

Setting $a = 1$ and $b = 0$, we see that

$$\sigma^2_{X+c} = \sigma^2_x = \sigma^2.$$

Corollary 4.8

Setting $b = 0$ and $c = 0$, we see that

$$\sigma_{aX}^2 = a^2 \sigma_X^2 = a^2 \sigma^2.$$

Corollary 4.9

If X and Y are independent random variables, then

$$\sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2.$$

Corollary 4.10

If X and Y are independent random variables, then

$$\sigma^2_{aX-bY} = a^2 \sigma^2_X + b^2 \sigma^2_Y.$$

Corollary 4.11

If $X_1, X_2, \ldots, X_n$ are independent random variables, then

$$\sigma^2_{a_1 X_1 + a_2 X_2 + \cdots + a_n X_n} = a_1^2 \sigma^2_{X_1} + a_2^2 \sigma^2_{X_2} + \cdots + a_n^2 \sigma^2_{X_n}.$$

Example 4.22

If X and Y are random variables with variances $\sigma_X^2 = 2$ and $\sigma_Y^2 = 4$ and covariance $\sigma_{XY} = -2$, find the variance of the random variable $Z = 3X - 4Y + 8$.

Section 4.4

Chebyshev's Theorem

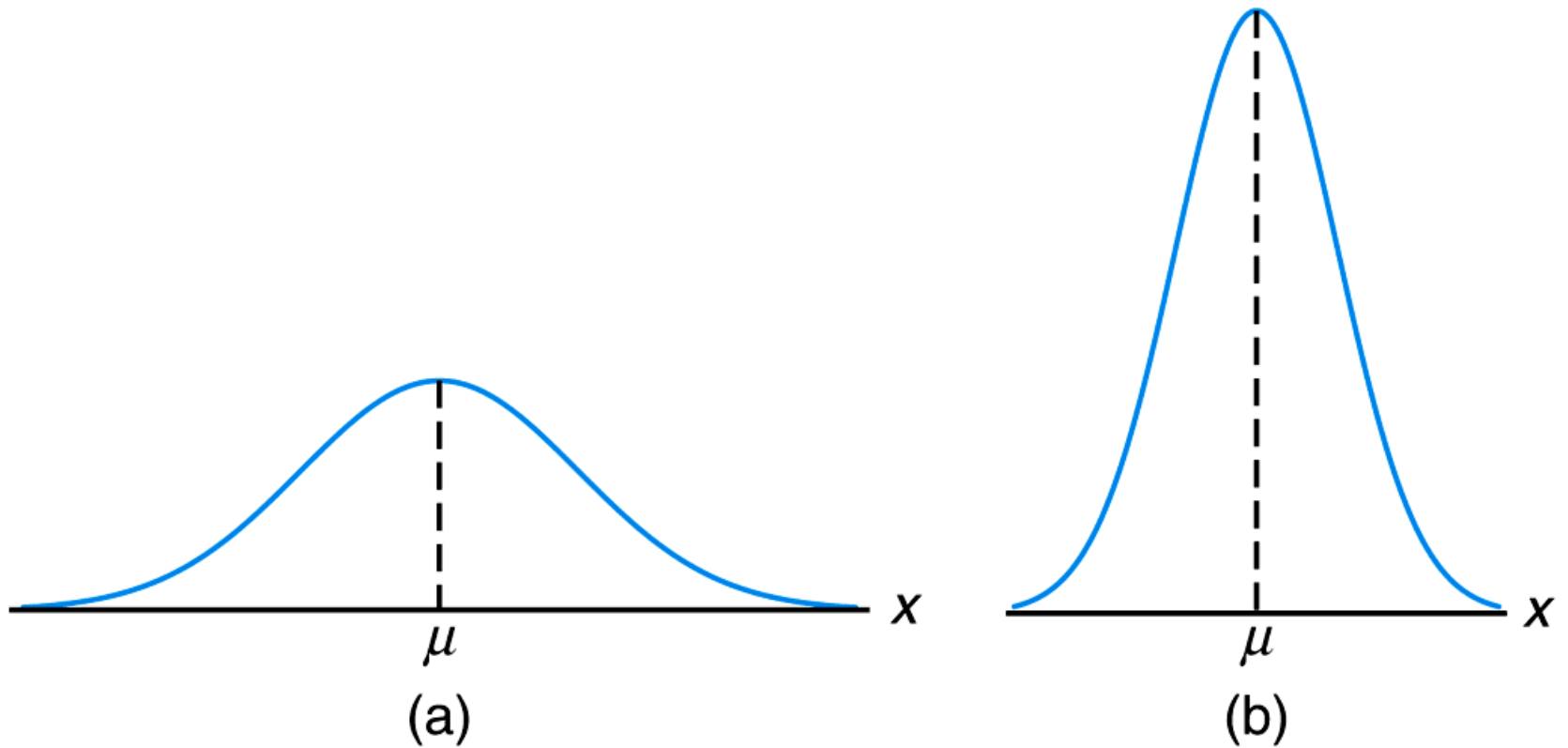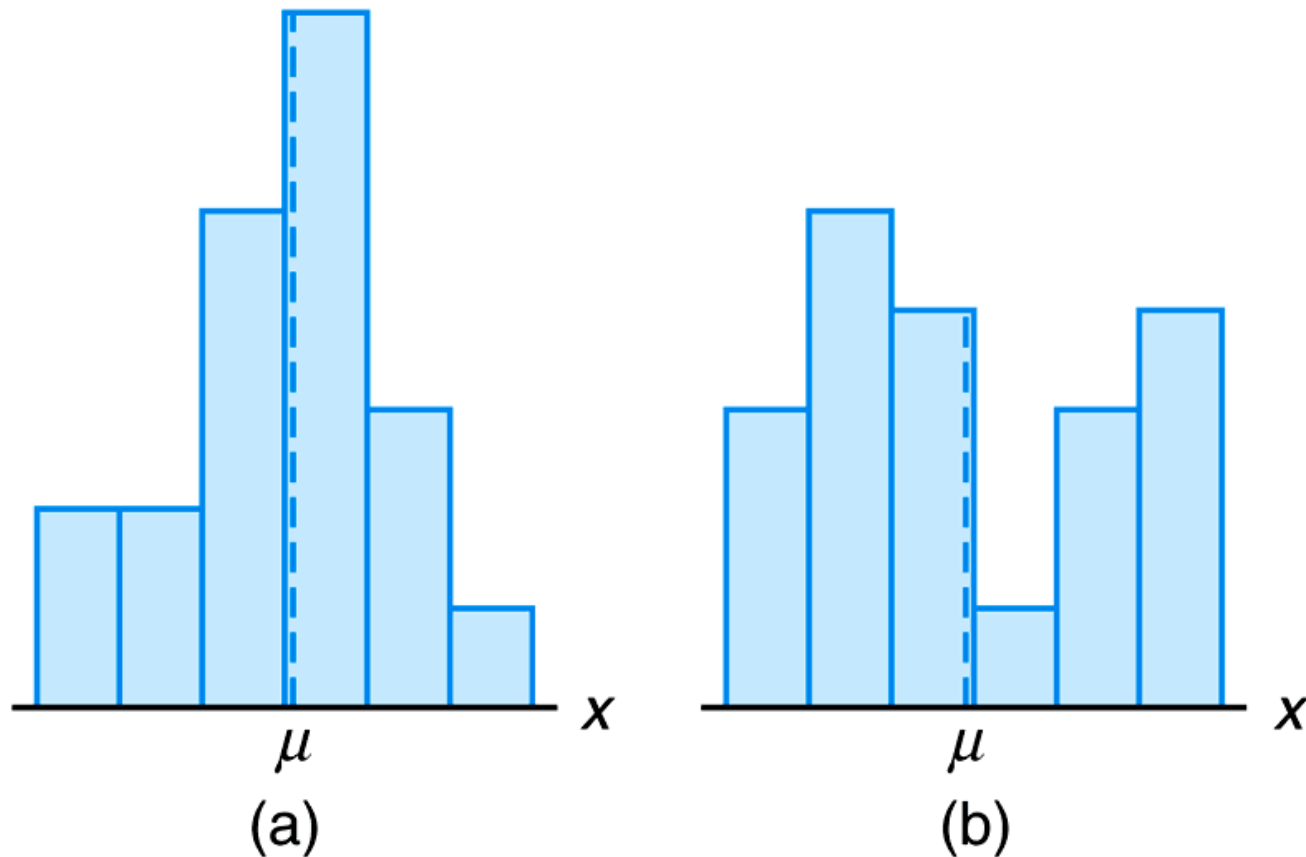**Figure 4.2** Variability of continuous observations about the mean

Figure 4.3 Variability of discrete observations about the mean

Theorem 4.10

Markov's inequality: Suppose X is nonnegative RV.

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

Theorem 4.10

(**Chebyshev's Theorem**) The probability that any random variable X will assume a value within k standard deviations of the mean is at least $1 - 1/k^2$. That is,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}.$$

apply Markov's ineq to (X-m)^2

Example 4.27

Example 4.27: A random variable X has a mean $\mu = 8$, a variance $\sigma^2 = 9$, and an unknown probability distribution. Find

(a) $P(-4 < X < 20)$,

(b) $P(|X - 8| \geq 6)$.